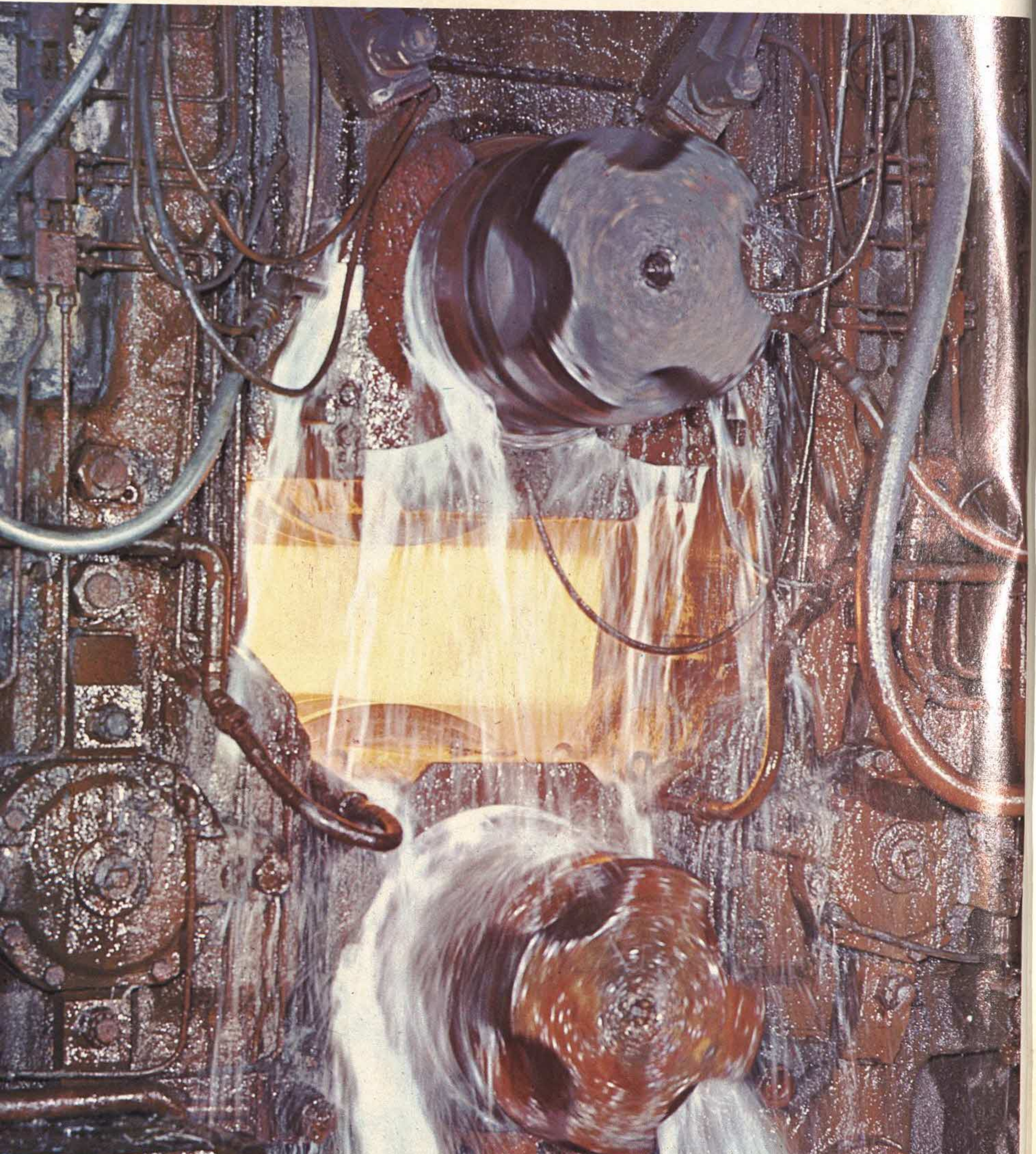
18

Illustrated World of Science Encyclopedia

Applied Science II:
MICROSCOPE, ELECTRON TO WIND TUNNEL

THE WORLD OF SCIENCE



ROLLING MILL—A white-hot steel ingot is flattened between water-cooled rollers.

THE WORLD OF SCIENCE

VOLUME
18
APPLIED SCIENCE II

Microscope, Electron, *to* Wind Tunnel

with

The Illustrated Science Dictionary

CREATIVE WORLD PUBLICATIONS, INC.
CHICAGO

SPECIAL CONSULTANT FOR VOLUME 18

Donald G. Lemke, Ph.D.
Assistant Professor of Mechanical Engineering,
University of Illinois at Chicago Circle,
Chicago, Illinois

Copyright © 1971 by CREATIVE WORLD PUBLICATIONS, INC.
Illustrations copyright © 1966 by Fratelli Fabbri Editori.

THE WORLD OF SCIENCE ENCYCLOPAEDIA has been translated and adapted from SCIENZA: ENCICLOPEDIA TECNICA E SCIENTIFICA by Fratelli Fabbri Editori, Milan, Italy.

THE ILLUSTRATED SCIENCE DICTIONARY, copyright © 1971 by F. E. Compton Co., appears with full permission of the publisher, F. E. Compton Co., Chicago, Illinois.

All rights—including but without limitation copyright of text, drawings or any part of them, and the copyright of translations into any language—are reserved to the copyright proprietors.

No part of this work may be reproduced or utilized in any form by any means, electronic or mechanical—including but without limitation photocopying, recording, or by any information storage and retrieval system—without permission in writing from the copyright proprietors.

Library of Congress Catalog Card Number: 71-139118

Typesetting primarily by SSPA Typesetting, Inc., Carmel, Indiana
Filmwork by Widen Engraving Co., Madison, Wisconsin
Printed and bound in U.S.A. by W. A. Krueger Co., Brookfield, Wisconsin

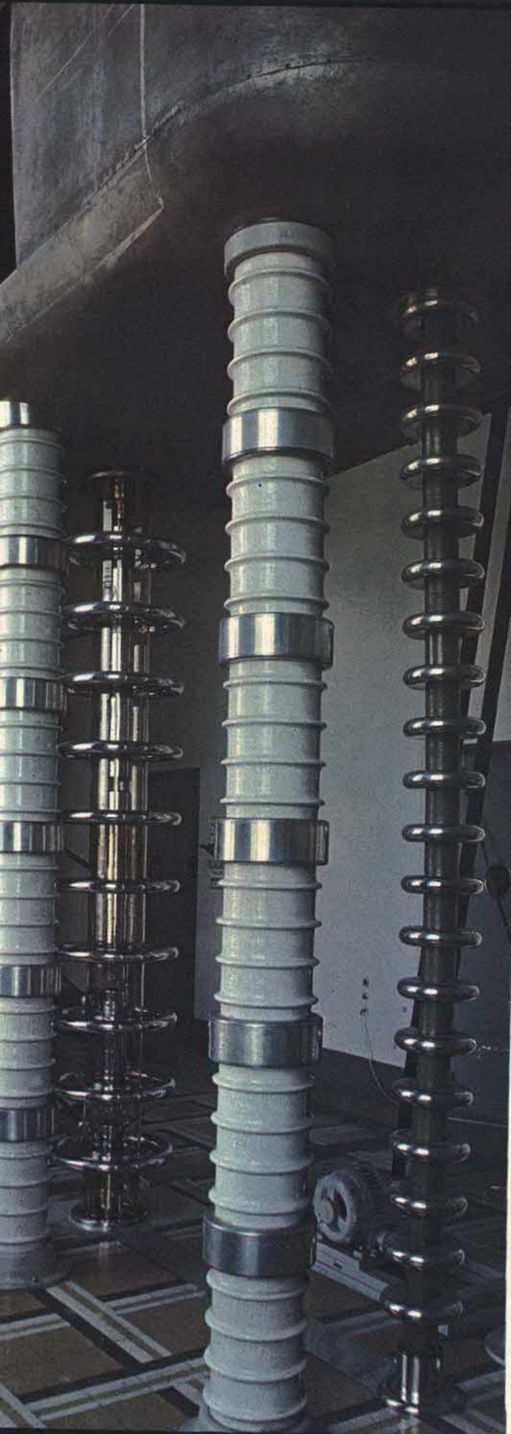
VOLUME 18

APPLIED SCIENCE II

Microscope, Electron *to* Wind Tunnel

APPLIED SCIENCE II

Microscope, Electron, *to* Wind Tunnel



INTRODUCTION

Today our society is experiencing grave misgivings about its past purchases in the name of technological improvement. Industrial pollution, inadequate waste disposal, chemical fertilizers, and insecticides all constitute real and major threats to the world's ecology. The demands for electrical power, pure water, clean air, and natural recreational areas have, in some locales, caused significant social, economic, and political problems. Antagonism has developed toward specific scientific programs, major types of industry, and technology in general. Just as yesterday's country store proprietor had to wait patiently for the child to choose his candy, today's technical community must cool its heels until the public decides. Will it be the moon and space exploration, nuclear power, environmental sciences, mass transportation, or medical research?

Just as technology provides increasingly deadly means of destruction, it also discloses principles that can be immensely beneficial. The knowledge and understanding of nuclear physics and radioactivity is a case in point. Discoveries in these areas have not only permitted man to create explosive devices capable of destroying cities and generating radioactivity levels that can mutate or destroy humanity, but have offered the promise of generating unlimited electrical power with a minimum of pollutants. Radiation therapy permits the elimination or arrest of certain critical ailments, and radiation chemistry provides compounds not possible by other means. These and other facts are treated in articles dealing with the nuclear explosion per se and with nuclear power. A third article deals with the particle accelerator.

Communication, by means of telephone and radio, is thoroughly represented in a series of articles dealing with the systems concepts utilized and the components composing these systems. Radar, a development of World War II, is included in conjunction with adequate air traffic control and weather warning communications. Remote control and monitoring of industrial processes is found in a discussion of telemetry and telesignaling; indeed, cooperation in the world community may be aided more by present and future achievements in this technical area than in any other now known.

Some idea of the expense and difficulty of resource development can be gained by reading the three articles dealing with the discovery, extraction, and processing of petroleum. Unless suitable synthetics can be developed, sources of this liquid will remain a matter of priority importance in the foreign policy of all nations.

Complementing the petroleum articles are articles on mining, on the Bessemer converter, and on steel processing and forging. Modernization and automation are the key words in the competition for the world steel market, and new techniques in these areas are shaking the economic world.

Photography is further explored in this volume in three articles that supplement the articles found in Volume 17 (camera). In addition, the use of ultraviolet light and the stroboscope is discussed.

Many more articles are included than those mentioned in this brief introduction. It is hoped that the articles presented herein will stimulate an expanded awareness, understanding, and dedicated concern for the implications of applied science with regard to humanity and its continuing existence.

DONALD G. LEMKE, Ph.D.,
Assistant Professor of Mechanical Engineering,
University of Illinois at Chicago Circle

THE MICROSCOPE—III | using electrons to provide magnified images

During the seventeenth century, when men looked through optical microscopes for the first time, a completely new world began to open before their eyes. Everything observed through the new instrument was seen greatly enlarged, revealing a complex structure that was invisible to the naked eye. Although the first microscopes were extremely primitive, they enabled men to study previously unknown phenomena, and they stimulated scientists to extend their field of observation and to construct better instruments.

Toward the end of the nineteenth century the optical microscope reached a high state of perfection. Since then relatively little improvement has been made in the clarity and the definition of microscopic images. The microscope has become the basic instrument of research in the fields of biology, medicine, and metallurgy; indeed, it has been used even for solving many problems in mechanics. Optical theory indicates that the construction of more powerful instruments using light rays and glass lenses is close to impossible. However, much evidence now exists to indicate that optical microscopes can reveal details that previously were regarded as submicroscopic. Optical microscopes can arrive at "seeing" particulars of the order of 250 Å—which corresponds to about one twentieth of the wavelength of light. Even so, many structures of interest to scientists are too small to be seen with a microscope that uses light rays.

In the early 1930s a group of German technicians constructed an electron microscope, an instrument that was capable of operating in a manner analogous to the optical microscope, but that utilized beams of electrons instead of rays of light. In place of the light bulb used in an optical microscope, the new instrument contained a source of electrons; in place of the glass lenses that serve to focus light rays and form the image in an optical microscope, the electron microscope contains magnets (called lenses by analogy) that serve to focus electron beams and form an image. With this system greatly magnified images of small objects can be constructed. The most important feature of the new system, however, is that these images are capable of revealing details that are much smaller than those that can be perceived with the optical microscope. Indeed, the electron microscope can show details that have diameters only 1/500 as large as those revealed by good optical microscopes.

Since the first electron microscopes were built, many improvements have been made, and researchers today possess an instrument that offers them extraordinary possibilities for exploring matter. Not only can they investigate the fine structures of cells, but they can perceive individual molecules and, with the help of special contrivances, even individual atoms.

Magnifications on the order of several hundred thousand times are necessary in order to observe such minute details, but these magnifications are possible because electrons behave like waves that are much shorter than waves of visible light.

THE LENSES OF OPTICAL AND ELECTRON MICROSCOPES

The electron microscope bears some resemblance to the optical microscope, because both instruments form magnified images by means of lenses. The lenses in the two instruments are quite different, however. Those of the optical microscope are usually made of glass. As light rays pass through the glass they are diffracted or bent (Illustration 2a). A properly ground lens forms a magnified image of the object under observation. Because glass cannot focus electron beams, another type of lens is necessary in the electron microscope. The electron lens is a circular magnet. As the electron beams pass through the opening in its center, they are focused by the magnetic field.

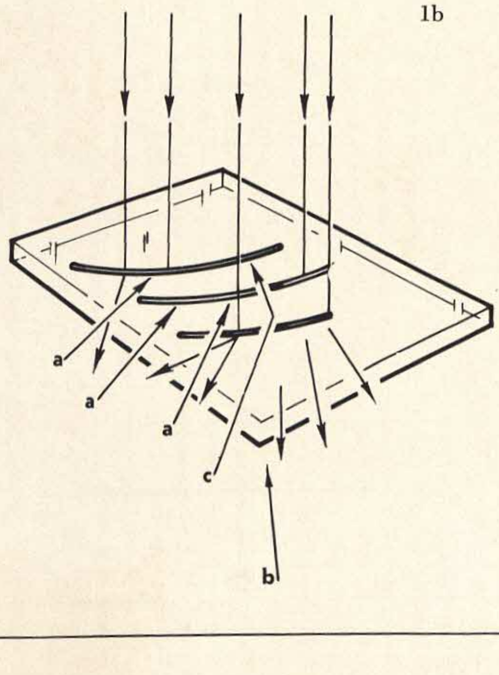
THE OPTICAL MICROSCOPE AND THE ELECTRON MICROSCOPE

The arrangements of lenses within the optical microscope (Illustration 4a) and the electron microscope (Illustration 4b) are similar. Each system has a condenser lens that directs beams of light or elec-

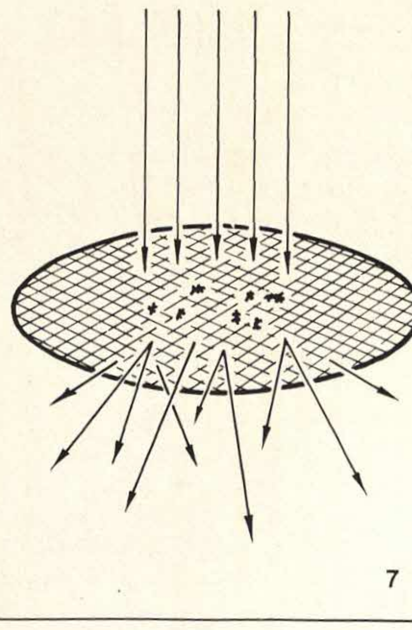
BEAMS OF LIGHT AND ELECTRONS—Illustration 1a is a diagram showing a beam of light passing through a microscope specimen. The specimen consists of thin fibers **a** placed between two glass plates **b.** Some light passes between one fiber and the next without being disturbed. Where it falls on the fibers **c,** some of it is scattered and some of it is absorbed by the fibers. When the specimen is observed from the side opposite the source of the light, the fibers are seen as dark-colored objects against a light background.

Illustration 1b is a diagram showing many electrons passing through a grid that contains some small crystals. Electrons are minute, subatomic particles, each of which carries a negative electric charge. Electrons can be accelerated by means of special devices and directed against a grid like the one shown in the diagram. When they strike it, some electrons pass through the meshes of the grid and between the crystals. Other electrons travel through the thinnest parts of the crystals and become scattered in all directions. If a screen covered with a layer of fluorescent material is placed behind this grid, the screen becomes illuminated where it is struck by electrons; it remains dark wherever the electrons have been absorbed. Therefore, what is seen on the screen is a "projection" of the form and the dimensions of the specimen. The object placed between the electron source and the screen absorbs electron rays in proportion to its own opaqueness to them. The most opaque parts of the object absorb the rays; the more transparent parts let them filter through. Specimens transparent to electrons or nearly so may be treated with electron-opaque materials.

1a



1b



trons toward the object. The objective lens of each microscope forms an image, which is further enlarged by another lens (the eyepiece in the optical microscope; the projector lens in the electron microscope). Because the electron beam is invisible to the human eye, the image is projected onto a fluorescent screen or a photographic film.

The projector lens of the electron microscope can magnify the image formed by the objective about 100 times. Because this image is already 100 times as great as the actual object, the final magnification is about 10,000 X. However, different magnifications can be obtained from both the objective and the projector. In fact, the images produced on the screens of electron microscopes can be from 4,000 to 200,000 X as great as the actual objects.

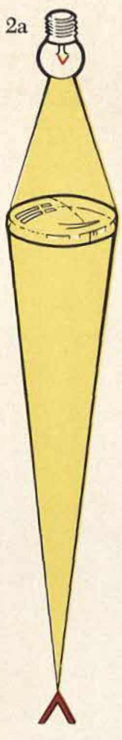
The images projected on the screen are extremely sharp and well defined if the instrument is perfectly adjusted. As a consequence, they can be photographed with films having very finely grained emulsions. The resulting photographs can then be enlarged further. Thus the final image may be magnified as many as a million times.

A good electron microscope provides images in which details with dimensions of less than 1 mm (about 0.04 in.) can be distinguished; these correspond to particulars of the object that are smaller than 0.000001 mm. This degree of enlargement and this sharpness of image permit scientists to perceive viruses, defects in crystalline structure, and chains of large molecules that form the most intimate structures of living tissues.

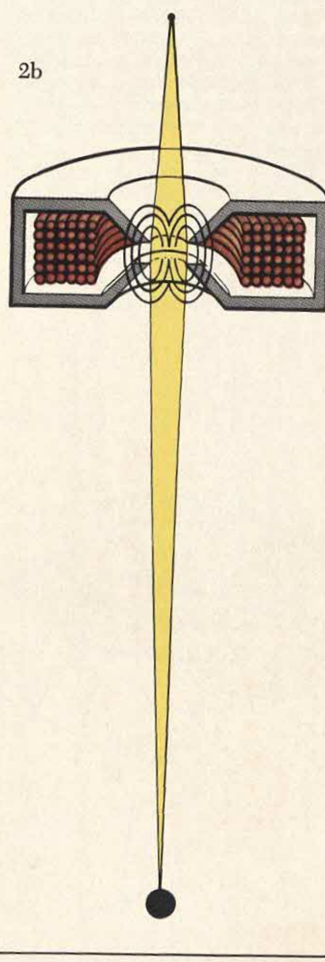
CONVERGING LENSES OF OPTICAL AND ELECTRON MICROSCOPES—A converging lens is a lens that can be used to enlarge images. If placed at an appropriate distance from a light bulb, it projects the image of the bulb onto a screen placed at a certain distance. The greater the ratio of the distance between the lens and the screen and the distance between the bulb and the lens, the greater is the magnification of the image. In the example presented in Illustration 2a, the image is magnified about three times.

The electron lens (Illustration 2b) is the equivalent of an optical lens. Essentially, it consists of a circular winding of wire carrying a direct electric current. This winding is enclosed in a box made of ferromagnetic material. When a beam of electrons passes through the central opening of the box, the electrons converge on a point after emerging from the lens. This point constitutes an "electronic image" of the object that emitted the electrons. Because the image cannot be seen by the human eye, it is projected onto a fluorescent screen. The screen is a surface covered with a special substance that becomes luminous when it is struck by electrons.

2a

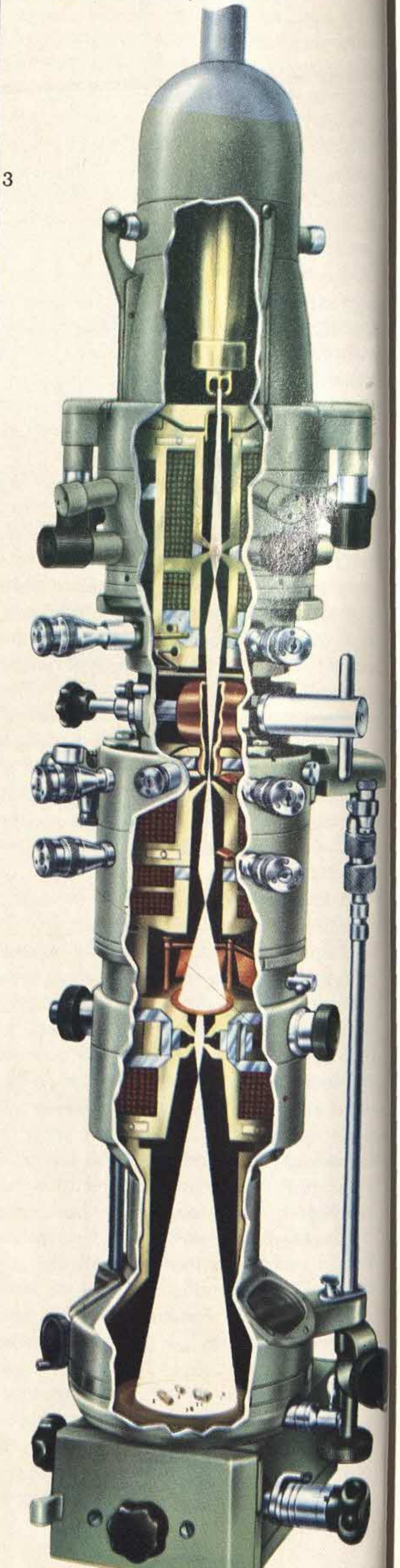


2b



CUTAWAY VIEW OF AN ELECTRON MICROSCOPE—This diagram shows the parts of an electron microscope and the path followed by the electron beam along the axis of the instrument. The electron beam is shown in the same manner as a beam of light in this illustration. In actuality, the electron microscope contains many other devices that serve to utilize the electron beam for carrying out observations of particular specimens.

3



4

SCHEMATIC DIAGRAMS OF OPTICAL AND ELECTRON MICROSCOPES—These diagrams serve to compare and contrast the structures and operating principles of optical and electron microscopes. The following list enumerates the various components of an optical microscope (Illustration 4a):

a. Light bulb.

b. Condensing lens, which gathers light rays emitted by the bulb and converges them on the specimen.

c. Specimen.

d. Object carrier, a small glass slide.

e. Objective consisting of several lenses.

f and g. Bundles of light rays that have passed through the specimen and form an image in front of the eyepiece.

h. Image of the microscopic specimen, magnified as a result of the projection of the objective.

i. Eyepiece, a magnifying lens that enlarges the image **h** formed by the objective (and that can serve to project this image onto a screen).

j and k. Bundles of light rays coming from the eyepiece and projected onto the plane **m.**

l. Image formed by the eyepiece.

m. Plane onto which the magnified image of the specimen is projected. (The image does not have to be viewed through the eyepiece; it can be projected onto a screen placed in the plane. Alternatively, a photographic plate or film may be used to record this image.)

Certain components of an electron microscope (Illustration 4b) are analogous to components of an optical microscope:

a'. Electron source, a tungsten filament that emits electrons when heated by an electric current. (The electrons pass through a number of electric magnets, which serve to accelerate the electrons as suitable voltages are applied.)

b'. Magnetic (condenser) lens that causes the electrons to converge on the specimen.

c'. Specimen.

d'. Grid supporting the specimen.

e'. Objective lens, which operates in a manner similar to that shown in Illustration 2b, collecting electrons that pass through **c'** and causing them to converge at **h'.**

f' and g'. Beams of electrons emerging from opposite sides of the specimen. (They are made to converge in opposite points of the image.)

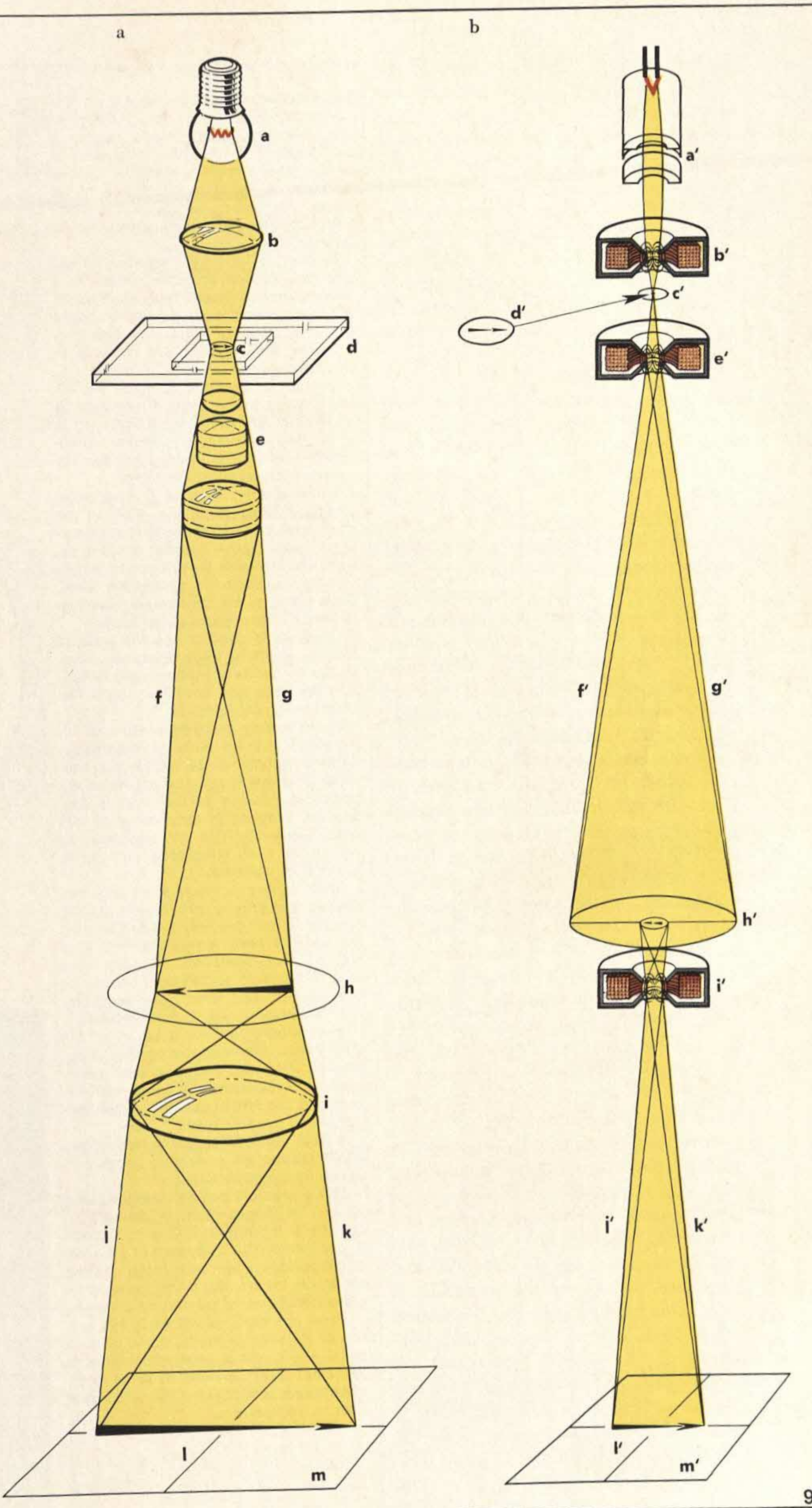
h'. First image of the object **c',** formed by the convergence of electrons on this plane. (This image, like the first image formed in an optical microscope, is further enlarged by another lens.)

i'. Projector lens that collects a part of the electrons coming from the image **h'** and causes them to converge into a second image **l'.** (The projector lens of the electron microscope serves the same function as the eyepiece in an optical microscope.)

j' and **k'.** Beams of electrons coming from two diametrically opposite points of the image **h'.** (They converge at two distinct points on the plane **m'.**)

l'. Second image.

m'. Plane on which the final image produced by the microscope lenses is formed. (A fluorescent screen is set up in this place and the electrons form a visible image on it.)



MINING | subsurface excavation

Mining is the industrial process of removing minerals from their place of natural occurrence. A mineral is almost any naturally occurring, nonliving substance found in the Earth, including metallic compounds, sand, oil, gas, and a host of other useful and valuable materials. Some of these are iron and copper used in the manufacture of automobiles, airplanes, and household appliances; salt for flavoring food; gold, silver, and diamonds for jewelry; uranium for atomic power; stone for building; phosphate to nourish plants; and gravel for driveways.

Man has mined the Earth for thousands of years. About 6,000 B.C. he dug pits and tunnels to obtain flint, a hard stone used to make tools and weapons. By 3,000 B.C. men were mining tin and copper, which they combined to make bronze, a hard alloy superior to flint, for shaping tools and weapons.

The ancient Romans recognized the wealth to be gained from mining and promptly took over the mines of every country they conquered. After the demise of the Roman Empire in the fifth century A.D., few advances were made in mining for the next thousand years.

Some mines, such as the Almadén mercury mines of Spain, the tin mines of Cornwall, England, and the zinc deposits of Silesia, have been worked for many centuries in Europe. Mining in North America did not begin until the period of colonization in the 1500s. Although the North American Indians used some metallic substances for weapons, paints, and utensils, they crudely fabricated them from minerals that they exploited only by ill-directed surface scavenging. Neither the Indians nor the Eskimos had even small mines, and they gave evidence of having had little, if any, use for gold, silver, and other precious metals.

In the eighteenth century, discovery of great iron and copper deposits in Minnesota and Michigan and the finding of zinc deposits in Missouri, Kansas, and Oklahoma, together with those of gold and silver in California, Nevada, and elsewhere, stimulated the evolution of modern metal mining in North America. Other stimuli were the growing population of the United States and the contemporaneous development of transcontinental railroads. More specifically, mining enterprises were spurred by the discovery of gold at Sutter's mill in California and later by the famous developments of Grass Valley and the Feather River; by those of the Comstock lode

1

EXCAVATING SHAFTS—Access to underground mines is by shaft, slope, or drift openings. Many mines employ a combination of these methods, and all mines have a minimum of two access openings to facilitate circulation of air through the mine and to provide alternative means of escape in case of emergency.

A shaft is a vertical opening driven through the rock from the surface to a mineral seam. Minerals from shaft mines must be hoisted to the surface in cages or skips. Complex machinery and controls and skilled labor are required in this operation.

Explosives are used to make the initial opening in the shaft. Illustration 1a shows how the charges are arranged in order to excavate a circular shaft. Charge holes are arranged differently to result in a rectangular shaft.

When working in loose ground, while sinking a deep shaft, the walls of the shaft must be lined to prevent their collapse. Some typical systems of lining the walls are shown in Illustration 1b, where wooden supports or props are used. Such supports are particularly effective in shafts with a rectangular section.

Concrete is used to line the walls of circular shafts of large diameters. Illustration 1c shows a section used to support the walls and constrain them. The concrete is cast in molds.

An extremely practical system of lining shafts of large diameter involves cylindrical segments of reinforced concrete (Illustration 1d). The segments are placed in position as the shaft is dug, and are attached to each other by iron bolts. Generally it is not necessary to wall in the shaft completely in order to prevent its collapse.

Shaft mining is much more selective than surface mining, with miners making an effort to remove only the ore, leaving the useless rock surrounding it. It is also a costly operation, because more time is required to follow the mass of ore. Copper, lead, zinc, gold, and silver are frequently mined by this method.

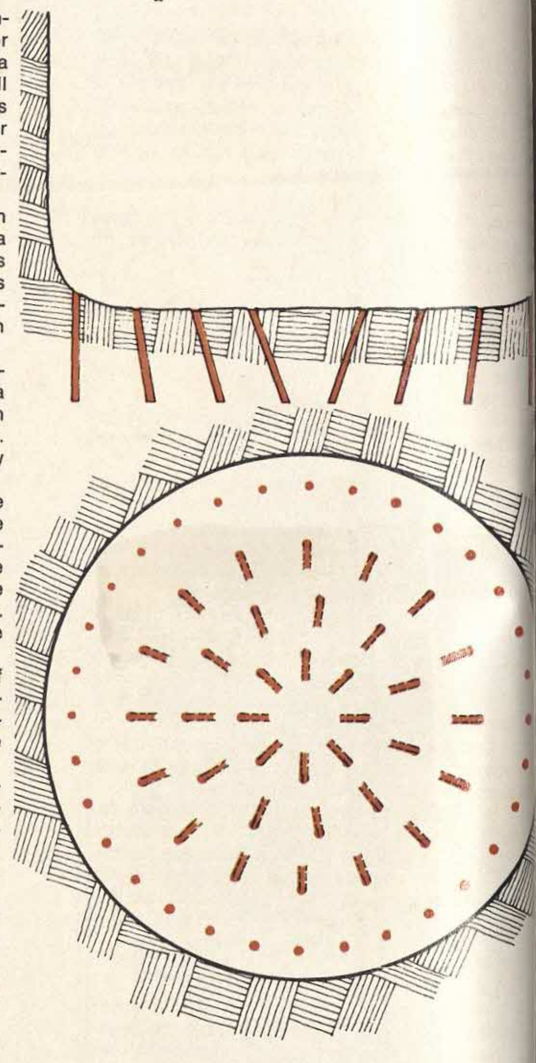
Illustration 1e shows a series of circular levels, each supporting special apparatus. After the shaft has been deepened by explosives, the loose material is loaded into large buckets that are then hoisted to ground level.

A large ventilation duct serves, above all, to remove the poisonous gases produced by the explosions.

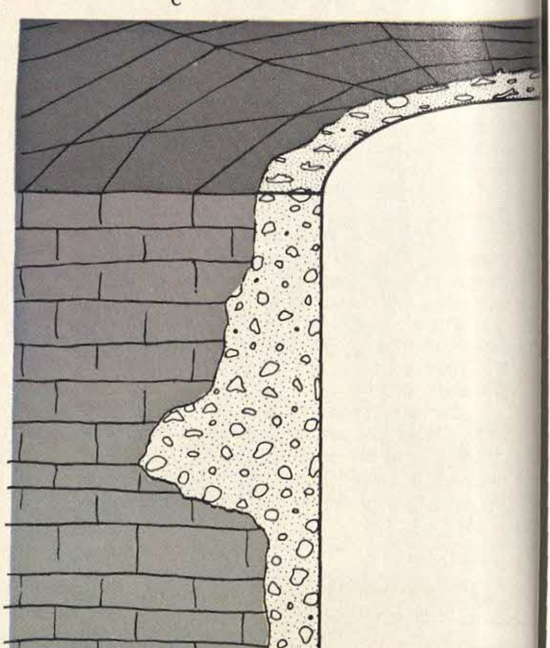
This equipment can be used for shafts up to 7 m (about 23 ft) in diameter. It can dig at the rate of 100 m (about 328 ft) per month. The lower part of the shaft is temporarily lined with metal. As the apparatus lowers, the metal lining is replaced with one of reinforced concrete.

There are many variations in this system of excavating shafts, some of them suitable to shafts of small diameter or to particular types of rock. In some methods the rock is perforated by drills rather than by explosives.

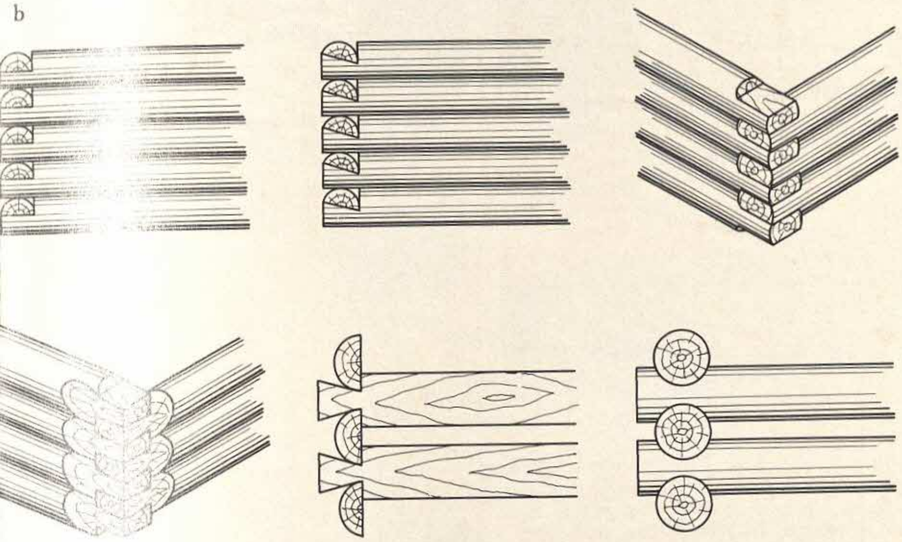
a



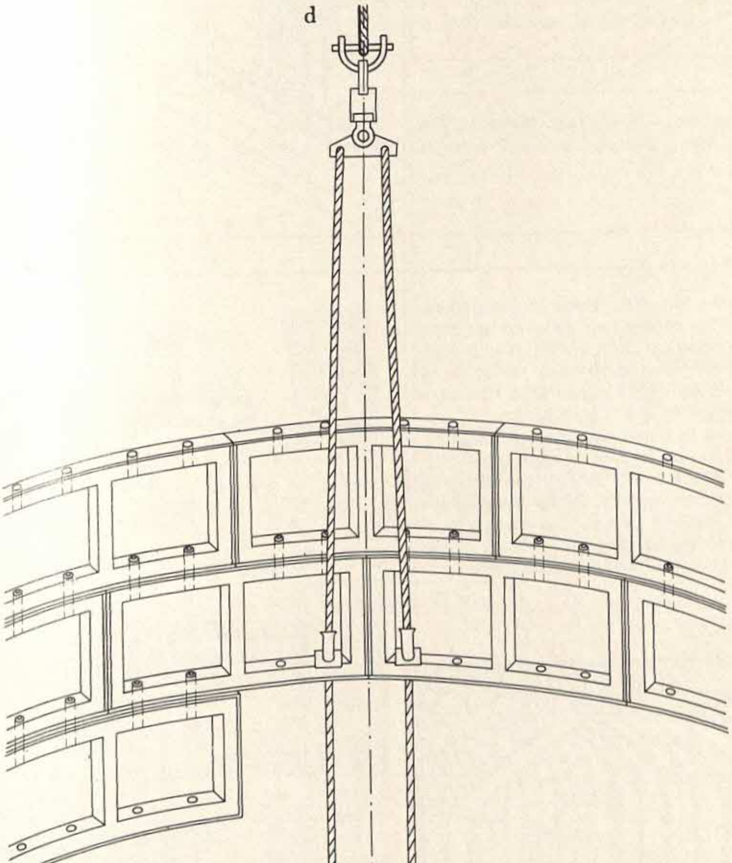
c



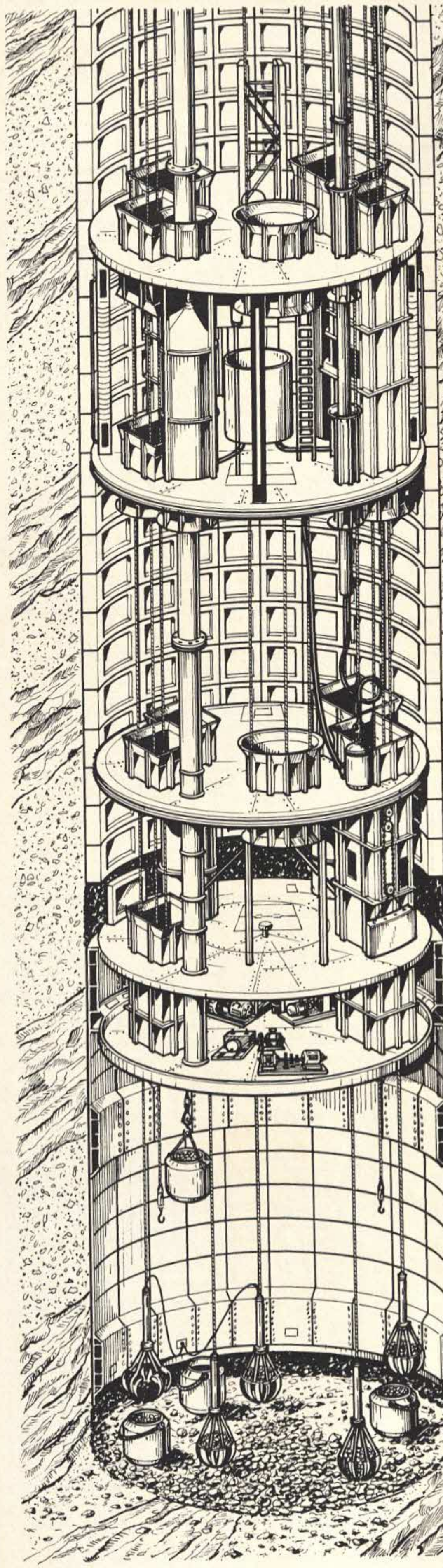
b



d

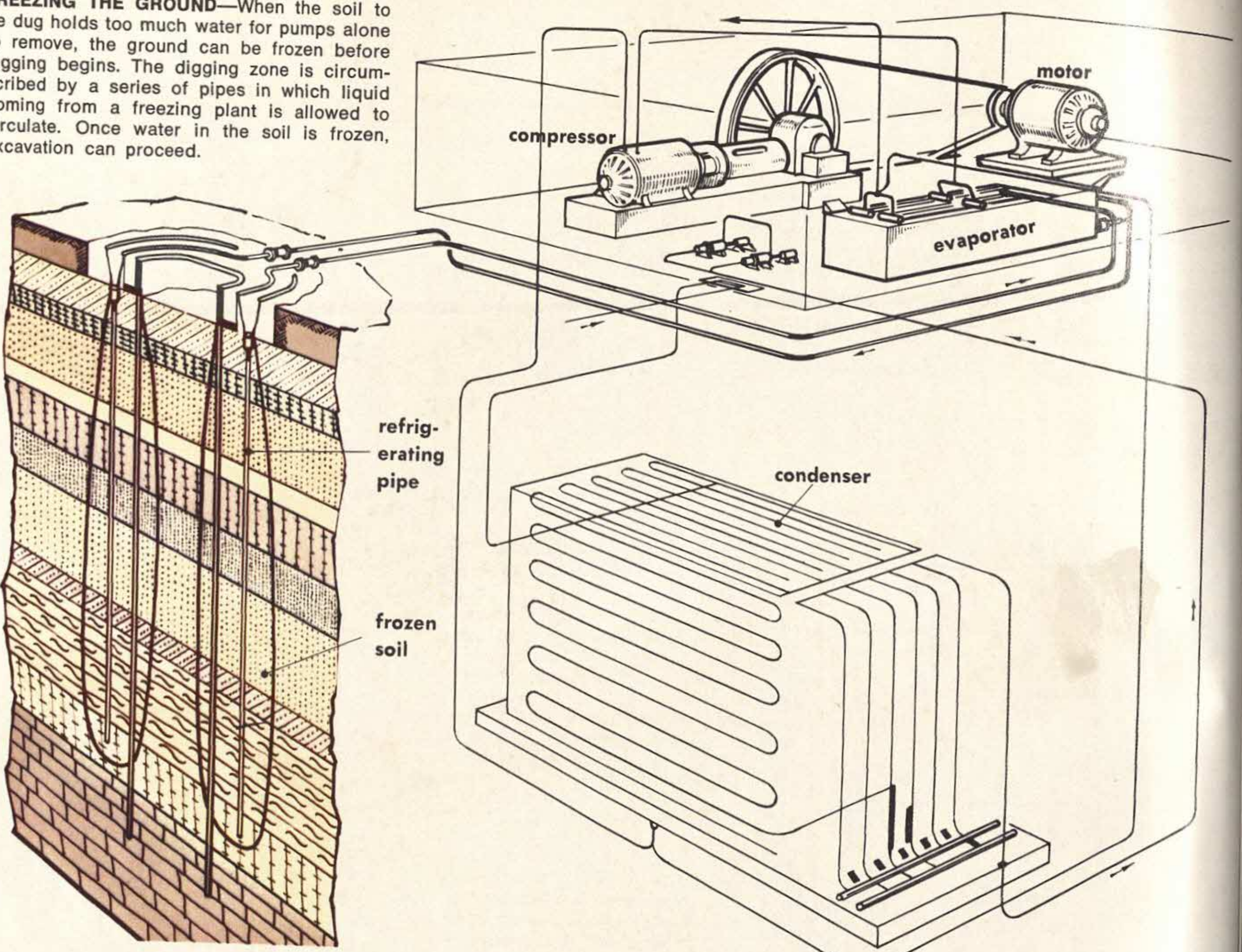


e



FREEZING THE GROUND—When the soil to be dug holds too much water for pumps alone to remove, the ground can be frozen before digging begins. The digging zone is circumscribed by a series of pipes in which liquid coming from a freezing plant is allowed to circulate. Once water in the soil is frozen, excavation can proceed.

2



3

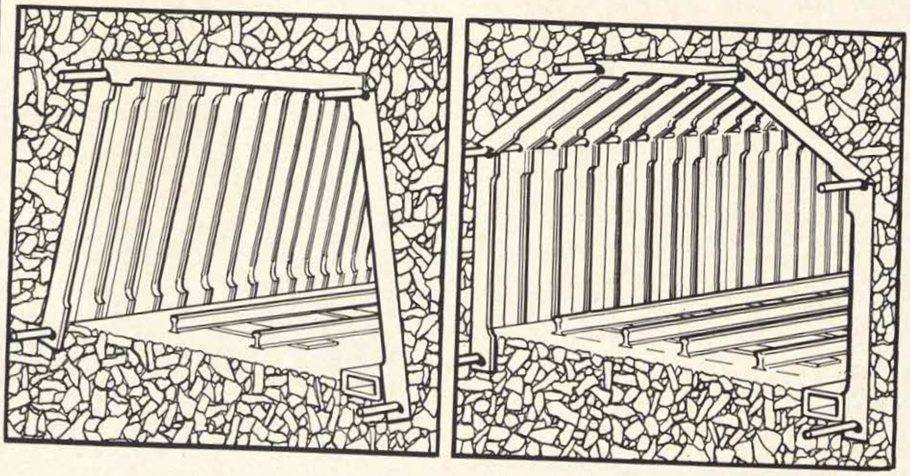
LINING GALLERIES—A gallery is a subterranean passageway that is essentially horizontal. Galleries in mines are reinforced with linings in order to prevent their collapse. Those galleries through which men and material pass are used throughout the mine and require more care than those used only during the extraction process.

Economically it is not feasible to use solid, lasting supports to deal with the slippage produced by extractions. A less costly method to stop reflex slips and to prevent the soil from sinking too much is to pile the gangue (the worthless rock from which the mineral has been extracted) into heaps in the gallery.

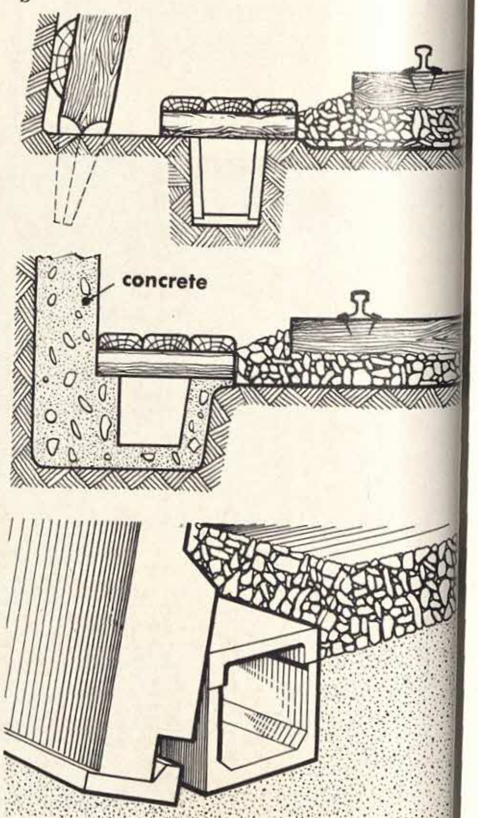
Illustration 3a shows two galleries used for transporting material. Both are lined with reinforced concrete. The strips are joined by a hinge to allow for slight slips in the roof or a slight movement in the supports.

Water, found to some extent in all galleries, is collected in an appropriate channel that carries it to a shaft from which it is removed by pumps. Illustration 3b shows three types of channels; the first is a simple one made of wood, the second is made of masonry, and the third is prefabricated from concrete.

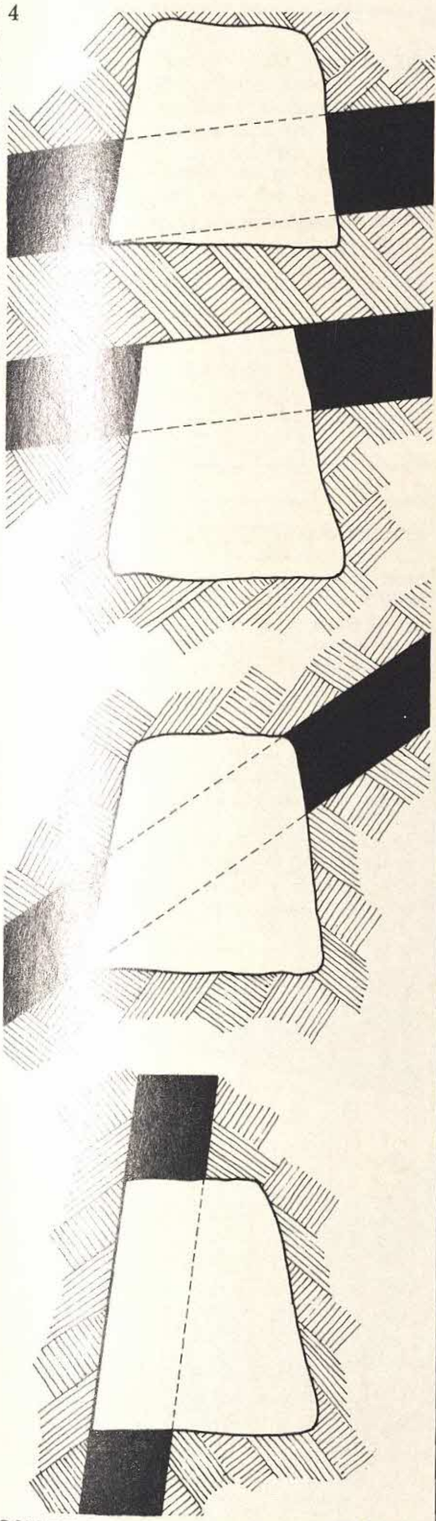
a



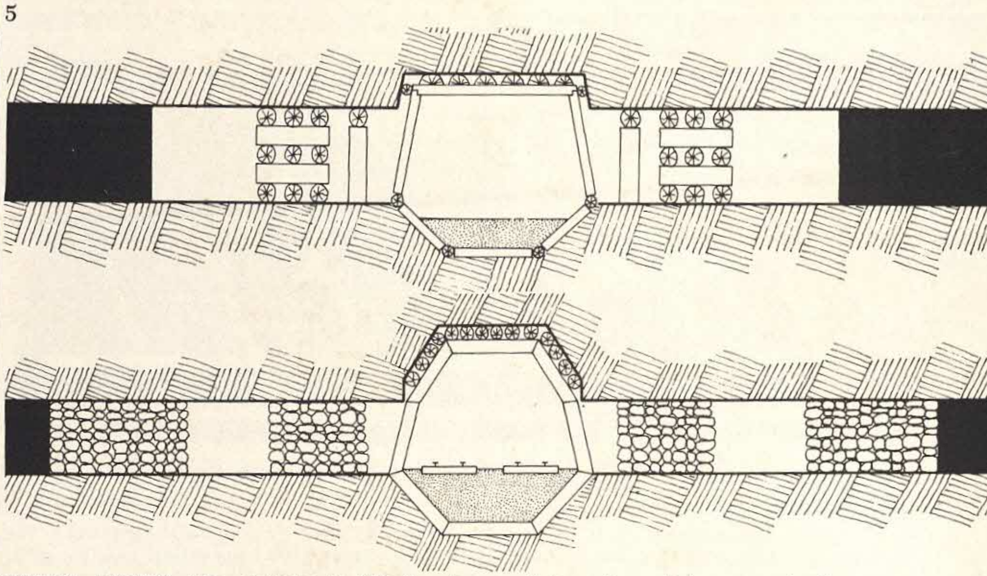
b



4



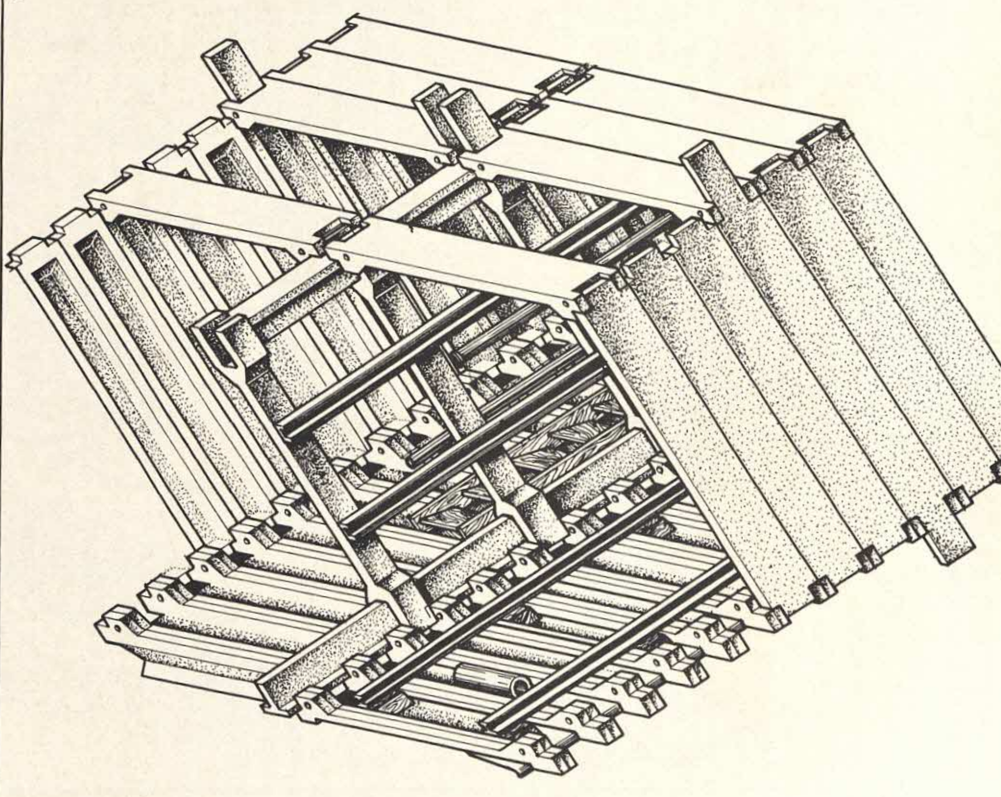
5



SPECIAL LININGS—The galleries in this illustration have been dug along a productive seam that is fairly long horizontally, creating a precarious equilibrium. To counteract this imbalance, the galleries require strong linings. At the same time, the excavated part of the seam must also be lined. If the excavation extends but a short distance, a wooden lining is sufficient. However, when the seam extends a long distance laterally away from the gallery, a lining of wood is economically unfeasible. In this case the hole left from the digging is filled with gangue. Today concrete is also used because of its strength and its relatively low cost. Iron is economical for large jobs and is especially effective in resisting great forces. Fiberglass, also used, has the same characteristics as wood, but has greater strength.

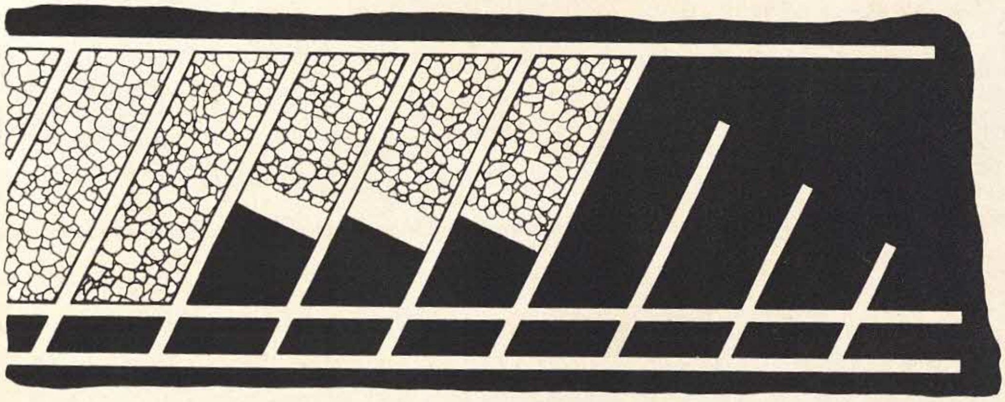
AN INCLINED GALLERY—Designed to accommodate wheeled vehicles, this gallery, with an angle of about 40°, is lined with concrete to prevent its collapse.

6



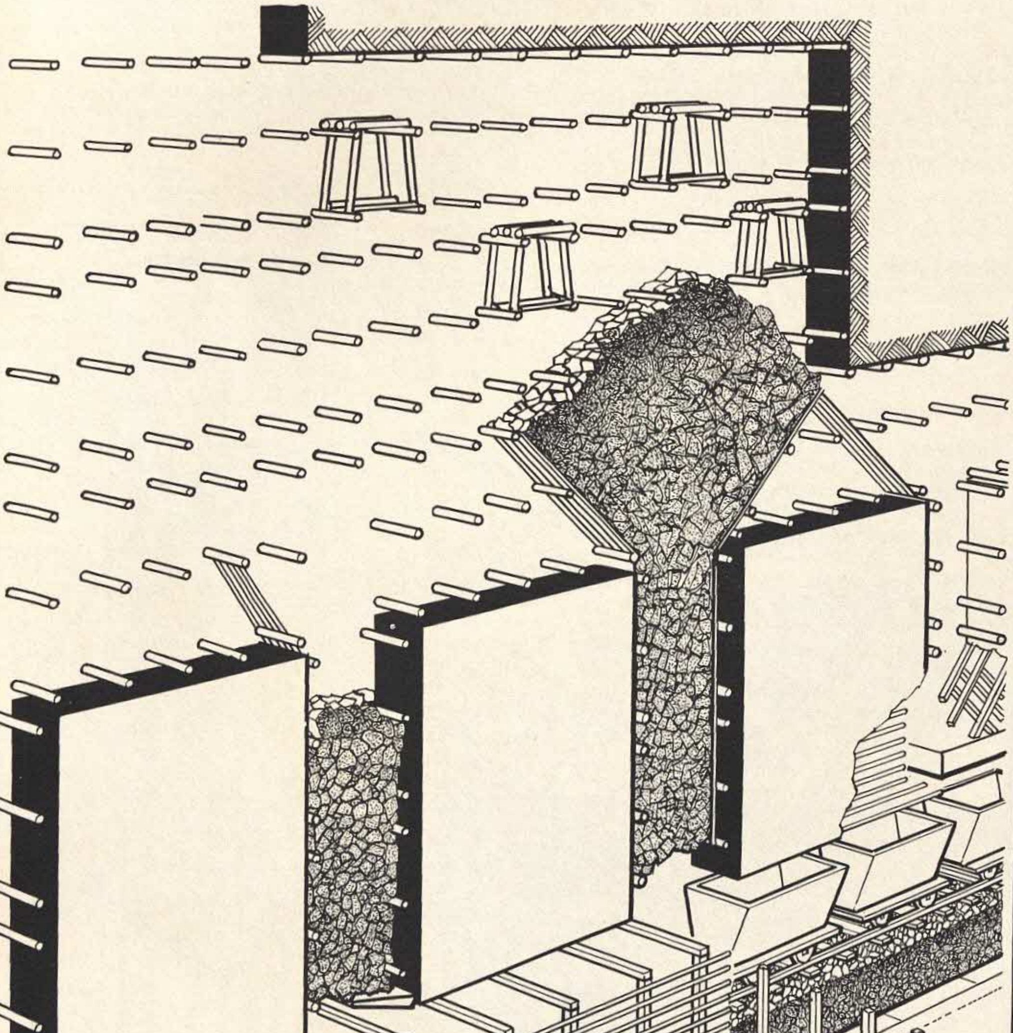
GALLERIES THROUGH AN EXCAVATED LEVEL—In order to pass through a thin, excavated level, a gallery must be opened with a section greater than that permitted by the thickness of the seam. The illustration shows the various angles at which the gallery can be opened in relation to the thickness of the seam. Each method requires a lining particularly suited to the method of opening. The galleries are made narrow at the top to achieve a balanced, arch-like form when the roof collapses. Galleries used only during the digging process require linings of only minimum strength.

7



EXCAVATING LONG SEAMS—When a seam or vein extends along a horizontal line, it is difficult to extract minerals without creating a rockfall. The illustration shows a method of opening a path through the mineral-bearing stratum. Two parallel galleries are linked by other galleries arranged diagonally in relation to the first two. The parallel galleries are extremely narrow and require no lining. The material between the two galleries is extracted and progressively replaced with worked rock that acts as a support. Included among the many advantages of using sterile rock (commonly called poor rock) as a filler are its low cost, the convenience of storing it belowground, and its use as a counterweight to hoist useful materials.

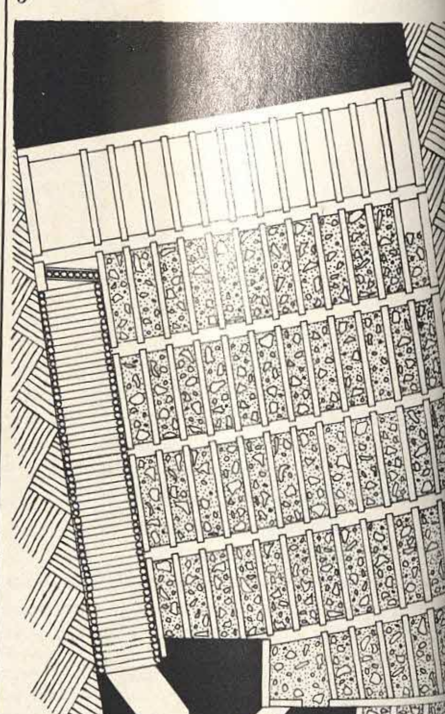
8 **EXTRACTION FROM VERTICAL SEAMS**—The technique used to extract material from a vertical seam is digging a gallery beneath the material and then raising shafts into it. The material falls from the shafts into the area still to be excavated, as shown in the illustration.



EXCAVATING EXTENSIVE MASSES—To excavate an extensive mass, adequate support must be provided for the walls of the cavity that is opened up. In the illustration an inclined seam of about 20 m (about 65 ft) has been dug out. Excavation of the mineral begins far above the gallery and is carried out in perpendicular blocks through the thickness of the seam. A row of blocks is removed and the hole is filled with poor rock. The mineral is sent through angled slopes to a lateral shaft kept open by a strong lining. From this shaft the material falls into trucks and is transported through the gallery.

Before excavation of any large mass, a careful study is conducted to determine if the material to be extracted is strong enough to support the openings of galleries.

9



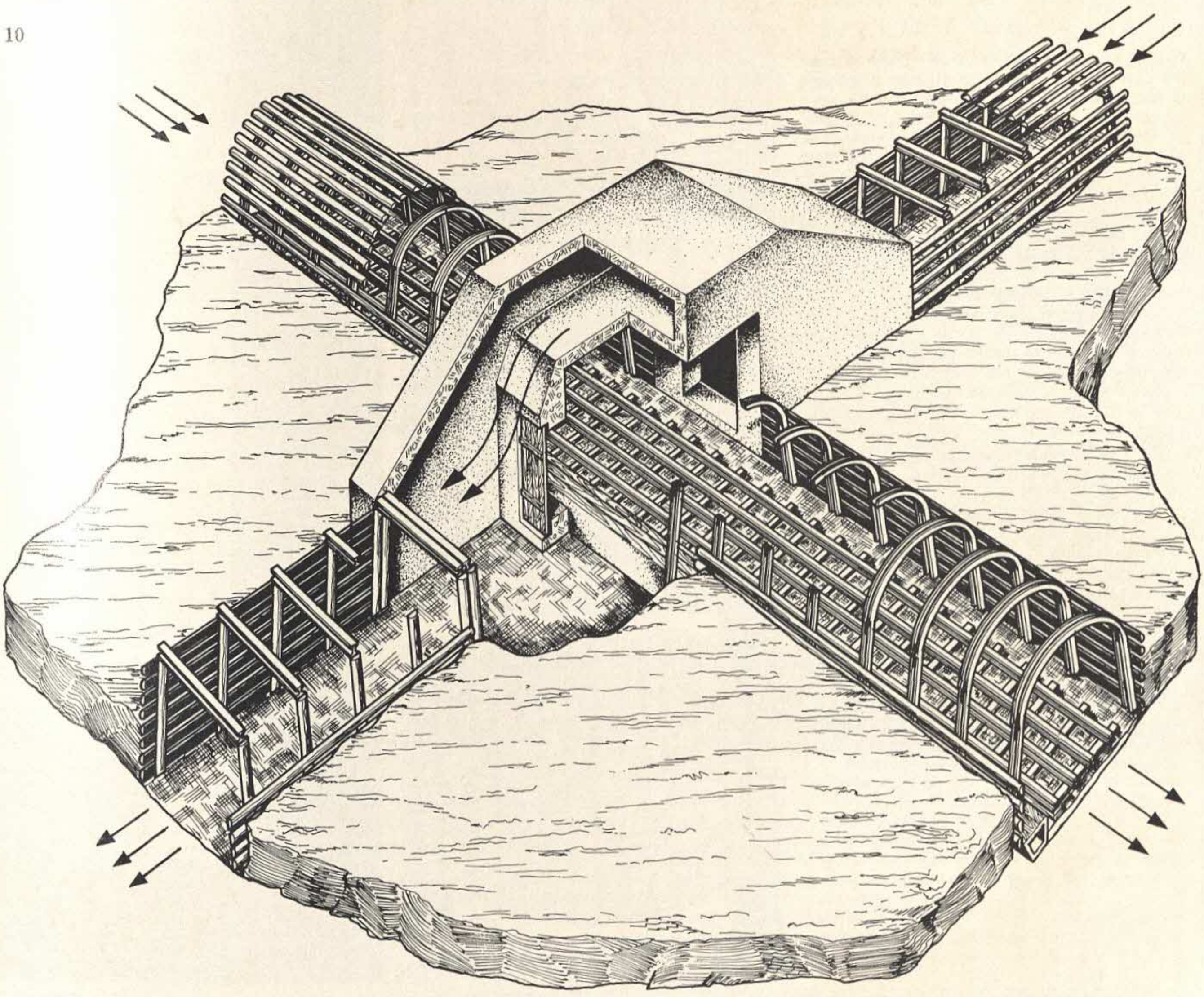
and Tonopah in Nevada; Cripple Creek, Leadville, and Silverton, Colorado; Tombstone, Jerome, Bisbee, and Morenci, Arizona; Virginia City, Helena, and Butte, Montana; Bingham, Park City, Ophir, and Eureka in Utah; Orofino, Bay Horse, and the Coeur d'Alene of Idaho; the Homestake of the Black Hills of South Dakota; the gold rush to the Yukon and to Alaska; and many other fabulously productive mining camps.

The search for mines was particularly intense after the Civil War, and much of the speed of the development of the western United States is attributable to it. The great wealth that accrued also served as an incentive to people of other lands, and great mines were discovered elsewhere in the world. Just as the Cornishmen of the tin-copper mines of England, for example, brought their skills to assist in the building of the great mines of the United States, American craftsmanship and ingenuity joined in the building of flourishing mining enterprises throughout the world. From Canada, Mexico, Venezuela, Brazil, Peru, Chile, Bolivia, South Africa, Rhodesia, Australia, Burma, India, and other countries, there has been a steadily growing production of an expanding range of metals. By the mid-twentieth century, the trend was to the increasing production of such comparatively uncommon metals as cobalt, titanium, nickel, and niobium (columbium), and a great mining rush to search out and develop deposits of uranium and related radioactive minerals swept much of the world.

The tendency in mining has been toward the increased use and improvement of mining machinery, so that modern mines are characterized by tremendous capacities. This has contributed to: (1) better working conditions; (2) the exploitation of lower-grade metal-bearing substances; and (3) the building of mines of great size and further promise.

PROBLEMS OF VENTILATION—Working underground involves the problem of adequate ventilation. The illustration shows the junction between two ventilation galleries. An overpass with thick reinforced concrete moldings ensures that the air between the two galleries will not mix. The two galleries must be kept separate because the purity of their air differs and the direction of air currents differs; the interaction of explosives detonated in each gallery must be prevented.

10



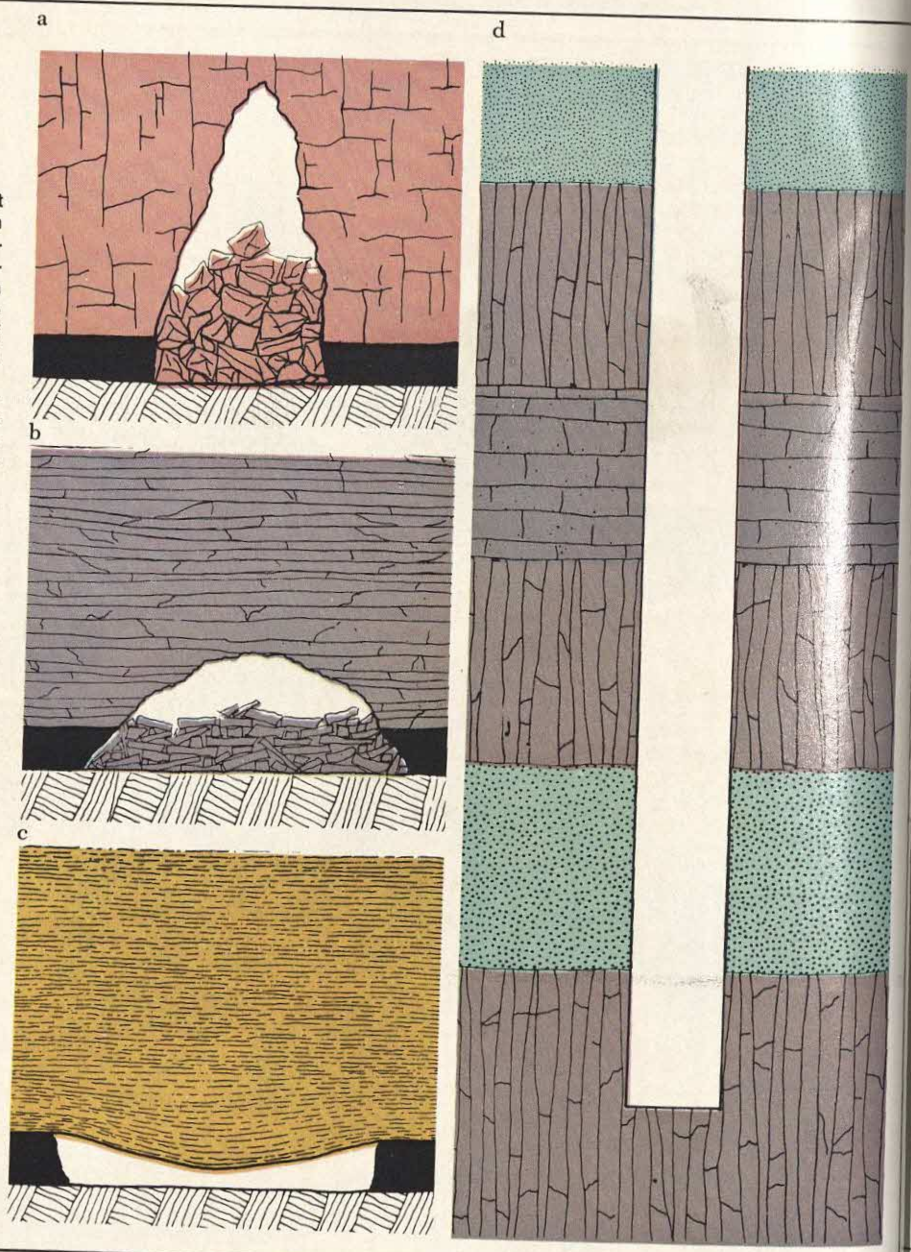
Many underground mines go to more than 5,000 ft in depth. Some noteworthy operations are conducted on the famous gold-bearing Witwatersrand of South Africa where some mine workings have reached depths exceeding 11,000 ft. As another example of the magnitude of some underground mines, a lead mine in southeast Missouri has excavations so extensive that they are served by two completely equipped underground machine shops of several acres in size. These shops are about 6 mi apart and are connected by mine passageways. All types of mining equipment used in the mines are repaired in these underground shops and some new devices of major size actually are built in them.

Because mining operations involve so many different types of machines and because they are frequently carried on in a subsurface environment where the responsibilities of the individual worker often are of highly abnormal scope, safety becomes of prime importance.

Although worldwide production figures are not available, the United States produces about half the world's total value.

Most mining ventures include the following operating procedures: (1) rock breaking (drilling and blasting), (2) mucking (loading), and (3) transportation (hauling and hoisting). Because of the wide variety of materials and because some operations take place on the surface and others underground, mining technique is very flexible in its operation. In this article some of the common methods of mining are examined.

11



WHEN THE ROOF CAVES IN—Imagine that the central part of a seam of coal has been emptied and left for some time without protection. In such a case the rock strata forming the roof will eventually fall in. The manner in which the roof collapses depends on the type of rock lying over the opening. The shape of the rock after it collapses will also depend on the type of rock involved.

A homogeneous rock that is split in a uniform fashion breaks into blocks, in some cases of large dimensions. These blocks fill the opening. The roof stops collapsing when it has assumed the shape of an arch (Illustration 11a). This shape enables the forces acting on the walls of the arch to remain in equilibrium with the resistance of the walls. In the case of hard rocks, the arch is quite pointed.

If a roof is made of shale in horizontal layers, the arch is less curved (Illustration 11b). In this case, too, the volume of material that falls from the roof is greater than the volume of material extracted.

With clay, the roof does not collapse to form an arch but simply changes shape, sinking until it almost totally fills the opening in the seam (Illustration 11c).

While it may be logically assumed that pressure in a vertical shaft (Illustration 11d) increases proportionally to depth, other factors also influence pressure. For example, a shaft passes through different kinds of rock, and the variations produce a pressure pattern that differs from that of depth alone. Another distorting factor is caused by water that impregnates rock and produces an increased pressure that cannot be precisely calculated.

MOTORS—I | two-pole electric circuits

The science that studies the effects of electric current, its transportation, and utilization is electricity. It deals with all the separate devices that, when combined, comprise electric circuits, electric networks, and electric power lines.

Electric circuits are conveyers of electric current; that is, electric charges move in them transporting energy from point to point, moving continuously inside conductors or wires.

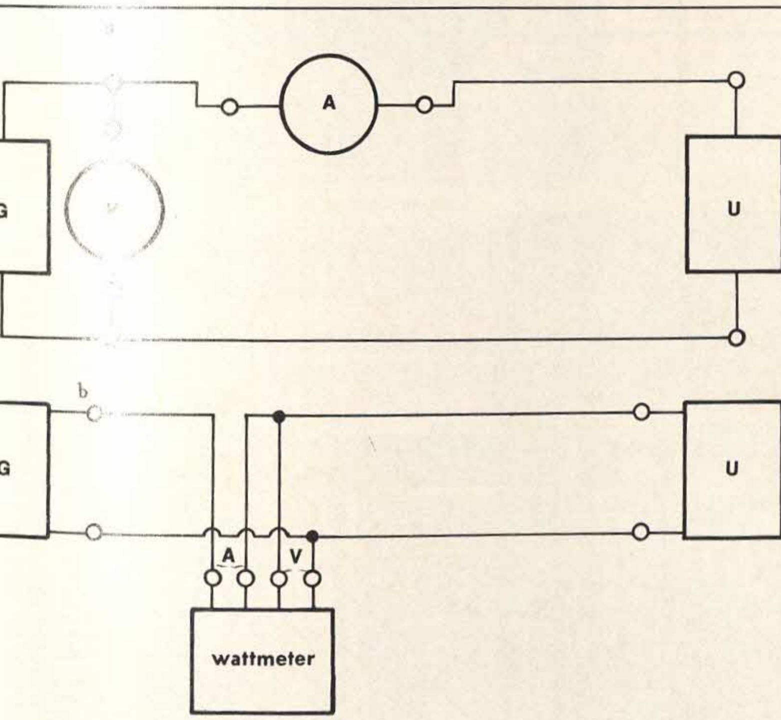
All circuits, the simplest to the most complicated, are made up of a set of fundamental elements called "dipoles." In electricity, the term "dipole" indicates a single element that has its own specific characteristics and properties. It is called a dipole because, inserted in a circuit, it is connected to the other elements by only two poles.

Circuits are physical systems with one degree of freedom; the electrical state is completely defined in them by two parameters: voltage (potential difference) and current. Once the voltage at either end of a dipole is fixed, the current that passes through can have only one constant value. From a study of the behavior of a dipole, a diagram depicting the values as a set of voltages can be drawn. The resulting curve in the diagram is called the "characteristic curve of the dipole"; it has an outline that, for that particular dipole, never changes unless the actual construction of the dipole is altered. The characteristic curve of the dipole is very important in electricity because if one of the two fundamental parameters is known, it is always possible to ascertain the other. Illustration 2 shows several characteristic curves of dipoles.

Electric circuits consist of a number of dipoles; two or more circuits connected to one another in various ways make up a network. In circuits and networks, dipoles can be connected to one another either directly or by a conductor. A conductor, too, is essentially a dipole because it has a certain resistance—although it is often negligible—and therefore acts like a resistor. A very long conductor is usually called a "line." The distinction between circuit and network is not generally respected. "Circuit," however, is the most common term and is very often used instead of "network."

Nearly all uses of electricity are also uses of power. This power may be 1,000 kilowatts for an electric locomotive, 100 watts for a lamp, or a fraction of a watt, such as a radio receiver picks up. Whether it is measured in kilowatts, in microwatts, or simply is read from the print on a lamp, the power and its basic unit, the watt, are central features in almost any point about the use of electricity even though the watt is not an electrical unit but is of a strictly mechanical nature.

1



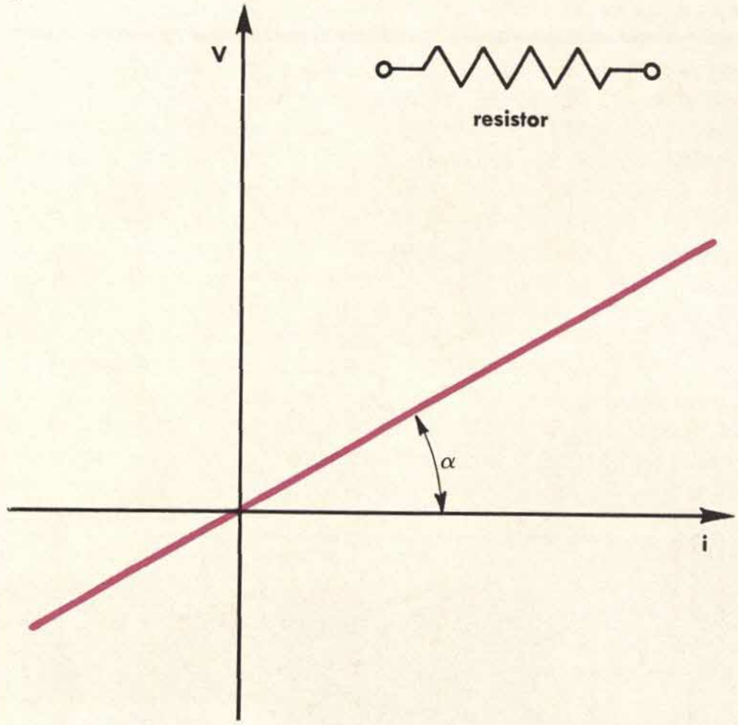
c

WORK AND ENERGY—The simplest type of elementary circuit is shown in Illustration 1a. In this circuit, **G** is a generator and **U** is a consumer. The current is measured by an ammeter **A** and the voltage by a voltmeter **V.** The generator is a dipole that transforms any form of energy into electric energy. The dynamo, for example, transforms mechanical energy into electric energy; the cells in Illustration 2d transform chemical energy into electric energy. When the generator is connected into the circuit, the generator supplies electric energy to it. In order to do this, however, the generator must perform a certain amount of work. The amount of work performed is proportional to the current, the voltage, and the length of time that current flows in the circuit. The consumers that use the energy supplied by the central station of a power company turn it into another kind of energy. A special device is used to measure the electric energy used in the home. It consists of an ammeter and voltmeter circuit (Illustration 1b). By relating current, voltage, and time, such a device can indicate the work performed by the generator, which is the same as saying the energy absorbed by the circuit. This device, shown in Illustration 1c, is a wattmeter. The unit used in measuring electric energy supplied to a consumer is the kilowatt-hour (kwhr).

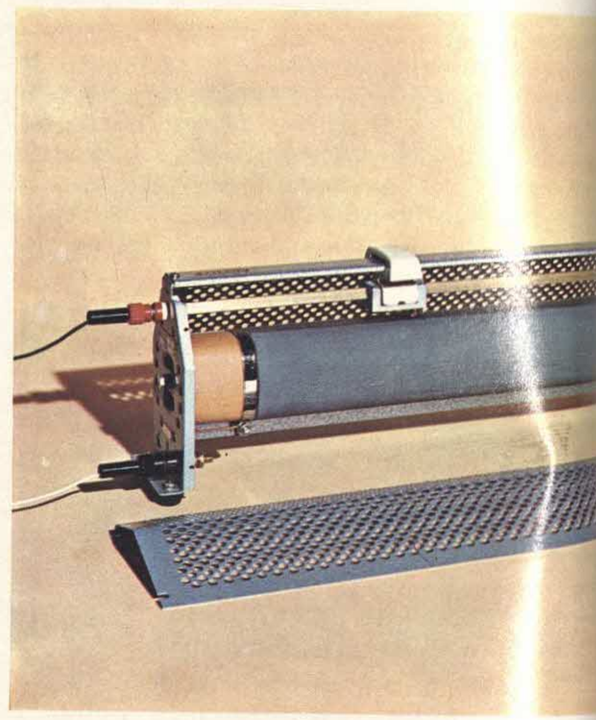
2

CHARACTERISTICS OF DIPOLES — The most common dipoles are shown in the illustration. The procedure for studying the behavior of any of them—when current is passing through them—is basically the same. Two instruments are necessary: an ammeter, to measure the current; and a voltmeter, to measure the potential. Once the current and potential are known, these may be plotted on a graph. After

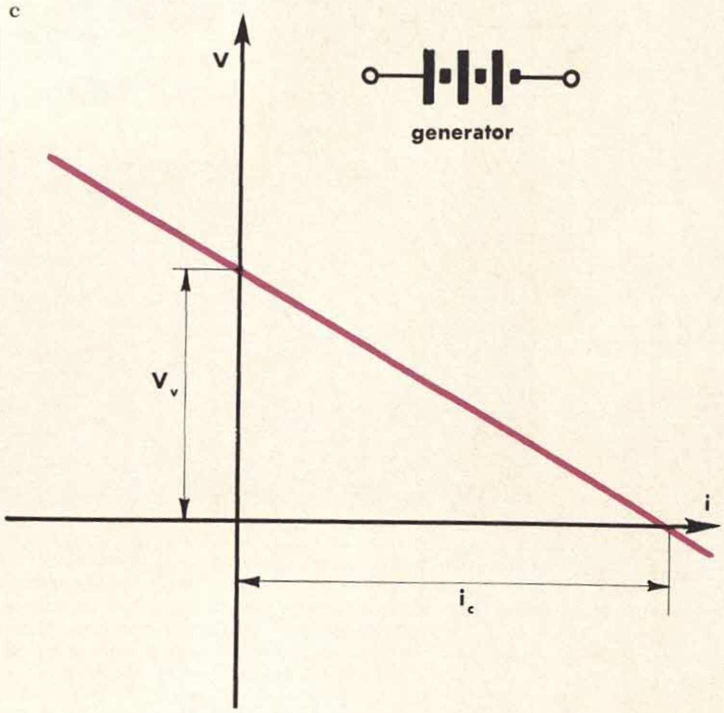
a



b



c



d

a certain number of measurements (at different values of current and potential), a curve can be drawn, connecting the plotted points representing the results. This curve is known as "the external characteristics of the dipole." From the shape of the curve, it is possible to classify dipoles.

One classification, for example, distinguishes dipoles with a rectilinear curve from those with curvilinear characteristics—the first are linear and the second nonlinear. Dipoles are also either inert or active, depending on whether or not the characteristic curve passes through the origin of the axes. For an inert dipole, when one of the parameters (potential or current) is zero, the other is zero also. This, however, is not true of active dipoles. Environmental conditions influence the shape of the curve significantly. In some cases, hysteresis is present.

Hysteresis phenomena are frequently met with in studying ferromagnetic materials. Another factor that can influence the shape of the characteristic curve is the heating to which the dipole is subjected as a result of the Joule effect. In experimentally determining the characteristic curve, it is necessary to take some precautions: voltage must be varied gradually in order to let the current stabilize at each change; also constant temperature conditions are necessary.

Illustration 2a shows the linear characteristic of a resistor—a dipole with a certain resistance. By varying the voltage **V**, the current varies linearly. Ohm's law is valid: $\mathbf{V = Ri}$. The resistance **R** gives the slope, or angle of inclination of the curve—that is, the tangent of the angle which the curve makes with the axis of the current **i**.

A rheostat is shown in Illustration 2b. This is a device that permits the value of the resistance to be varied. If the measurements are repeated with different values of the resistance, the characteristic curve is still a straight line, but its slope differs from the previous one.

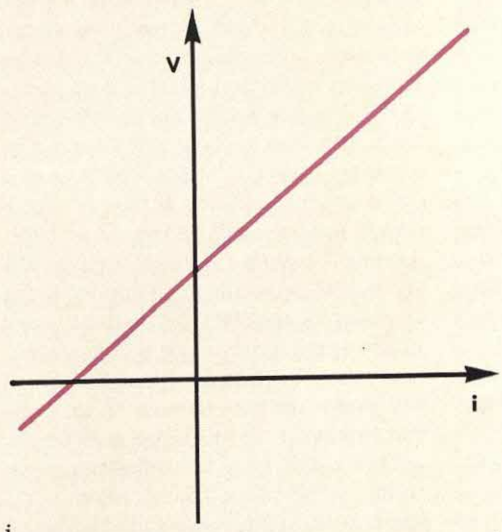
Illustration 2c shows the characteristic curve of a generator, such as the cells shown in Illustration 2d. Here, the curve is rectilinear, but it does not pass through the origin of the axes. When voltage is zero, a current $\mathbf{i_c}$ flows through the generator; this current is called "short-circuit current." When no current flows, a difference in potential can still be measured at the poles of the generator. This voltage is called "no-load voltage" $\mathbf{V_v}$.

A consumer is any dipole that uses the energy supplied by a current, whether it is used thermally, mechanically, or in any other way. A common "consumer" of energy is a simple electric motor. Illustration 2e shows the characteristic curve of a consumer such as the electric motor shown in Illustration 2f.

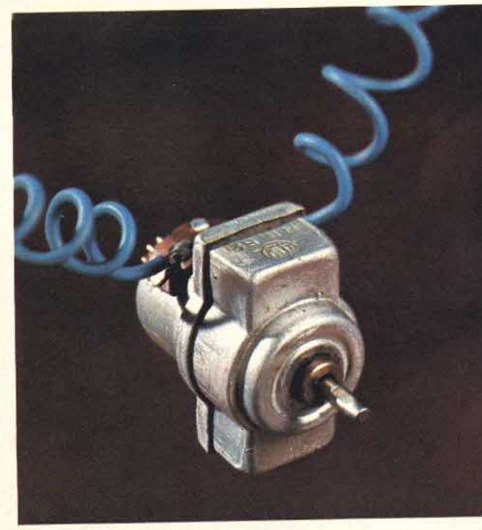
Illustrations 2g and 2h show the characteristic curves of a normal light bulb (2g) and an arc lamp (2h); an arc lamp is shown in Illustration 2i. The variations in the curve of the arc lamp, under different conditions of voltage and current, are clearly distinguished in Illustration 2h.

Illustration 2j shows the curve of a rectifier.

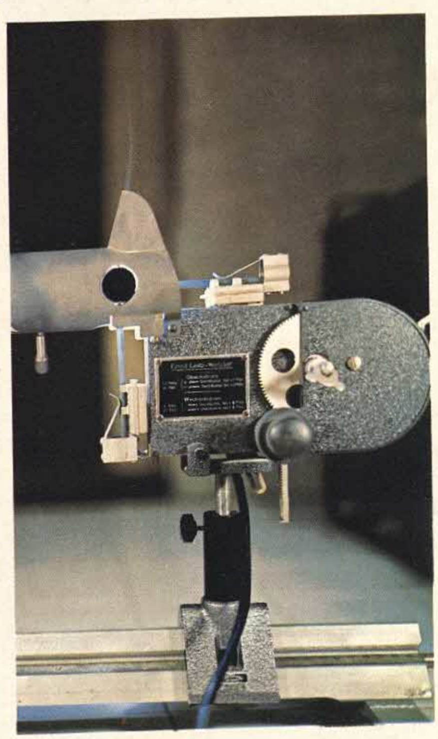
e



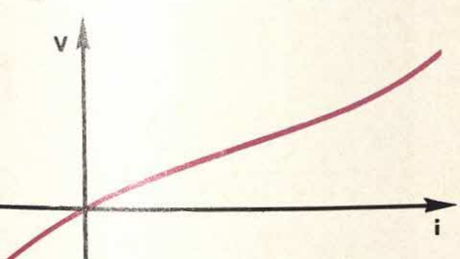
f



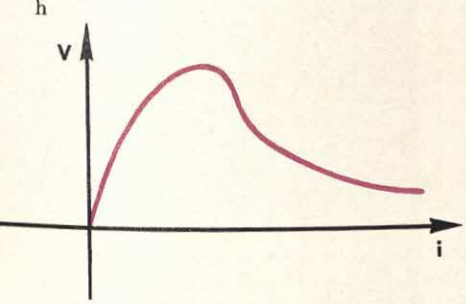
i



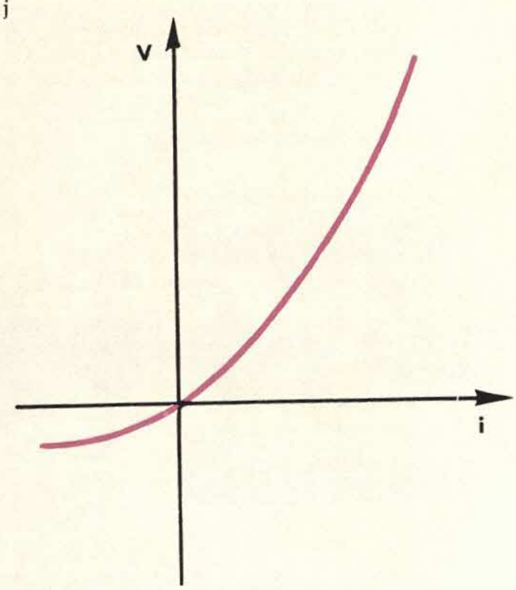
g



h



j



MOTORS—II | principles of jet propulsion

The progress made in the field of aeronautics during the last twenty years has been so great that now each new conquest tends to go almost unobserved, despite the fact that such developments would have been considered as science fiction less than half a century ago. Even in the field of astronautics (which might be defined as aeronautics beyond the limits of the terrestrial atmosphere) each new launching only arouses a fraction of the interest and astonishment that accompanied the launching of the first Sputnik in 1957. The lunar landings are almost the only occasions for widespread interest and excitement. This enormous progress has been made possible by the rediscovery of jet propulsion. The term "rediscovery" is used because the rocket, which is only one application of jet propulsion, has been known to the West since medieval times and the Chinese were using it in the eighth century for both warfare and fireworks.

Until the beginning of the present century the rocket remained relegated to relatively minor roles as a propellant of fireworks or as a means of signaling. Its use in warfare was practically abandoned upon the advent of firearms, which were far more efficient and effective.

The principle underlying jet propulsion was disclosed as a result of the broadening that the study of physics and, more particularly, dynamics had received from the English scientist and mathematician Sir Isaac Newton. The advent of modern technology enabled this principle to be translated into practical terms.

JET PROPULSION AND JET ENGINES

The principle of jet propulsion is basically simple and can be easily illustrated with some examples. Connect a garden hose to a faucet and attach a hose nozzle to the other end (the nozzle greatly reduces the cross section of the issuing jet of water). Now lay the hose on the ground and turn on the faucet. If the pressure is strong, the hose will be pushed violently backward and may even continue to jerk about. This phenomenon is applied to the operation of certain types of rotary lawn sprayers that are commonly used for watering large fields and gardens. The thrust that causes the hose to move and the one that causes the forward motion of jet aircraft or a rocket are two different aspects of this same phenomenon. Another example is the action of a balloon when it is inflated and allowed to escape from the hands.

In general, when a mass of any kind (particularly a fluid) is rapidly expelled, the object from which it has been expelled (the hose or balloon, for example) becomes subject to a thrust that causes it to move in the direction opposite to that of the motion of the mass. This principle can be applied to both rockets and jet engines.

A rocket is simply a pipe closed at one end (it takes in no ambient air) that burns fuel and ejects the products of combustion through an exhaust nozzle, and is thereby projected in the opposite direction. The German V-2 rockets of World War II are examples of this type as were the ancient Chinese rockets, although on a much cruder scale. The propulsive components of a rocket (fuel, reagents, and associated gear) are referred to as rocket engines.

A jet engine varies in several ways from this closed pipe design and operation. One basic difference is that it includes the intake of air in its operation. It derives its velocity from air intake, fuel combustion, and ejection of combustion products at the nozzle (rear) end.

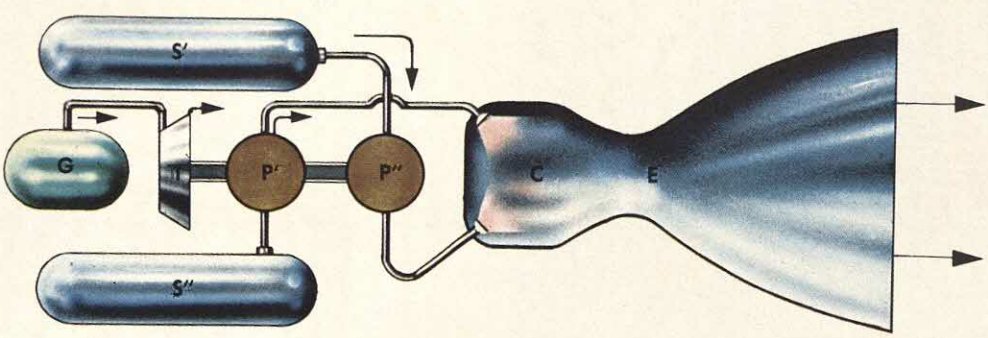
One type of jet propulsion, then, strictly speaking, is nothing more than a rocket, descended from the fireworks of the Chinese, in which the mass of the propel-

1a

G: gas generator to drive auxiliary turbine

T: auxiliary turbine

P′ P″: propellant pumps



E: exhaust nozzle (De Laval type)

S′ S″: propellant tanks

C: combustion chamber

1b

lant (which includes the oxidizer) is carried by the vehicle itself. The other type involves intake of air as the oxidizer does, although the same basic principle applies to both the rocket and jet engine.

In jet propulsion, the expulsion of a mass of fluid (usually gas) from a vehicle subjects the vehicle to a thrust F. This thrust is approximately given by the formula $F = mW/t$, where m is the mass of expelled fluid, W the speed of the fluid, and t the time during which expulsion occurs.

In both types of engines, it is possible to impress the necessary speed on the mass with either the help of mechanical means (the screw employed in water propulsion, for example), or by exploiting the heat energy produced by combustion (or by electrical or magnetic energy). Heat energy, however, is the means applied to jet engines by using the heat energy produced by a combustion inside the engine itself. In fact, even a ship moves forward by virtue of the thrust developed as a result of the water displacement produced by the screw. However, the screw receives its motion of rotation from a completely different type of engine.

Simple types of jet engines are the ramjet and the pulse-jet, which (except for a valve system in the pulse-jet) have no moving parts. Other more complex and hybrid types are the turbojets, turbofans, turboprops (propjets), turborockets, turboramjets, and so forth. In the ramjet, the exit speed that generates the thrust is ensured by the high pressure that is generated inside the combustion chamber as the result of the chemical reactions that produce large quantities of heat. The turbojet is also derived from this principle. Pressure in the ramjet is obtained by means of an air inlet—a tube shaped so that a part of the kinetic energy possessed by the air in its motion relative to the vehicle becomes transformed into pressure. In the pulse-jet an intermittent combustion occurs in closed or semiclosed chambers; the increase of the temperature and the volume of the gaseous mass causes the pressure increase. Various combinations of operations are possible—in the turbojets, for example, the air arrives at a compressor after it has already been compressed by an air inlet.

ROCKET ENGINES AND PROPELLANTS

To date, the rocket engine is the only chemical propulsive device capable of powered flight in a vacuum. Rocket engine design is simple and consists of two major elements—the combustion chamber and the nozzle. Many types of rockets employ a converging-diverging device called a De Laval nozzle (from the Swedish inventor Carl de Laval), in which the exit speed reaches values of several times the speed of sound and generates the thrust needed to drive the vehicle. Necessary pressure is generated by exothermic chemical reactions—particularly by the combustion of a propellant—all of which are onboard. The propellant can be solid, liquid, or mixed. A solid propellant occupies the whole of the combustion chamber and the principal problem consists of controlling flame propagation and maintaining combustion stability. In the case of a liquid propellant, on the other hand, the critical design consideration concerns two liquids, the fuel and the combustion supporter or oxidizer. These are introduced into the combustion chamber by means of an injection system. The pressure for injection is generated by carefully controlled pumps.

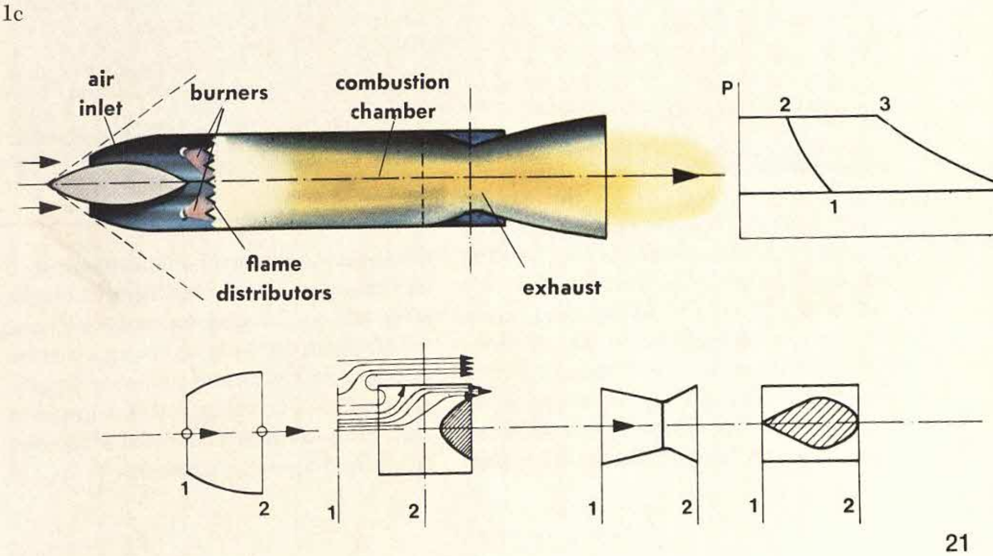
Depending on the type of ignition required, liquid propellants are classified as hypergolic or nonhypergolic. A hypergolic propellant ignites spontaneously when its fuel and oxidizer are brought together. Nonhypergolic propellants are

ROCKET ENGINE AND RAMJET—Several types of rockets exist, but the only type employed today for both astronautical and military uses is the thermal rocket. This rocket exploits the large amounts of thermal energy produced by chemical reactions between the propellant materials. Illustration 1a shows the general layout of a liquid propellant rocket engine. The two liquids, one of which is a fuel and the other an oxidizer, are contained in two tanks **S** and **S′**, which generally occupy the whole or almost the whole of the structure of the rocket. The two liquids that constitute the propellant are introduced into the combustion chamber by means of two pumps that are operated by an auxiliary gas turbine. The mass m of gas leaving the nozzle at a speed W during a period of time t generates a thrust F, given by $F = mW/t$. The gas generator **G** drives the auxiliary turbine **T**, which operates the propellant pumps **P′** and **P″**. The nozzle **E**, behind the combustion chamber **C**, is a De Laval type. The exit of the gases is violent, almost explosive (Illustration 1b).

Illustration 1c is a schematic representation of a ramjet engine. An extremely simple jet engine, it does not require any machinery other than a fuel-supply pump. The ramjet engine consists of an air inlet, a combustion chamber, and an exhaust nozzle. The air inlet is a special configuration that makes it possible to obtain a pressure increase in the current of air that enters the suction tube. It is shaped in such a way that the cross-sectional area increases when the relative speed of the intake air with respect to the aircraft is less than the speed of sound. Conversely, when the relative speed is greater than the speed of sound, the cross-sectional area must first decrease and then increase. An engine of this type can be put into operation only when the vehicle already possesses a sufficient speed of its own. The compressed air passes from the inlet into the combustion chamber, which contains the fuel burners. These burners are generally followed by a flame-stabilizing device.

The thermodynamic cycle shown in the right portion of the illustration is identical to that of the gas turbine.

1c



ignited electrically. The substances required to sustain the reaction of propellants are known as oxidizers. Liquid oxygen, nitric acid, fluorine, and strong solutions of hydrogen peroxide are oxidizing agents. Hydrogen, the hydrocarbons, aniline, and hydrazine, on the other hand, are fuels (the last two fuels are hypergolic with respect to nitric acid).

Solid propellants are mixtures of nitrocellulose and nitroglycerine, or other mixtures consisting of an oxidizing agent and a fuel.

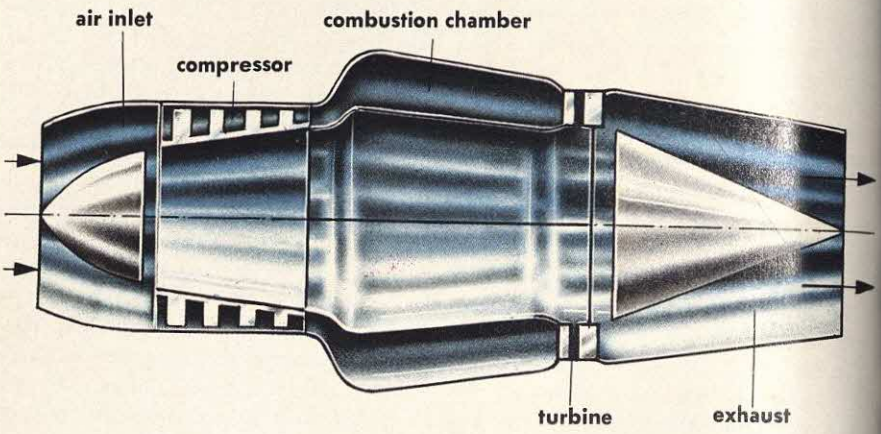
THE TURBOJET AND TURBOFAN

The turbojet and turbofan are the most widely used jet engines. They have now almost completely replaced the old propeller and turboprop engines. Turbojets were applied for the first time to some military aircraft during World War II and have since been rapidly improved and perfected. Their introduction immediately led to the shattering of old speed limits—not only of military aircraft but of commercial aircraft. The distances between one continent and another no longer seem as great as they did a few years ago, while the number of passengers and the quantity of goods that can be carried annually continue to multiply rapidly.

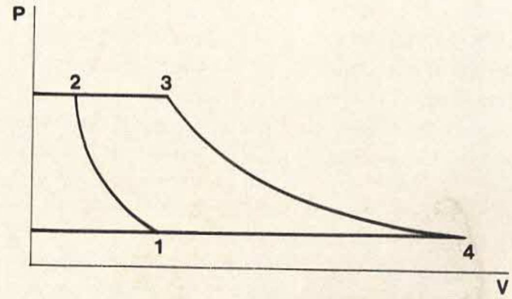
The turbojet engine employs an air inlet that produces a preliminary compression of the mass of air that is sucked in. This air is then further compressed by an axial compressor prior to passing into one or more combustion chambers in which a measured quantity of fuel is being burned. Consequently, a considerable increase in the volume and the pressure of the mass of the gas occurs; this gas is allowed to expand through a turbine where it acquires speed. The speed continues to increase in the exhaust tube (tail pipe) until it reaches the value necessary to create the required thrust. The turbine through which the gas expands operates the air compressor. The thermodynamic cycle is substantially the same as that of the gas turbine.

The airstream of a conventional turbojet speeds rearward faster than the airplane moves forward, causing turbulence and loss of thrust. The turbofan engine (also called fanjet) incorporates a fan that directs some of the intake air around the engine. The remainder of the air goes into the combustion chambers. Both airflows join at the rear to create a larger, slower jet stream, thus increasing engine efficiency at low speeds and reducing exhaust noise. The turbofan takes in about four times as much air as the turbojet and is half again as efficient.

2a



2b



THE TURBOJET—The turbojet is the type of jet engine most widely used. The fundamental parts of this engine are shown in the illustration. The external air enters through an inlet that performs the same function in the ramjet, and ensures a preliminary compression of the air. This compression effect is increased considerably during flight because of the aircraft's speed. From the air inlet, the air passes into the compressor, which is usually a multistage compressor of the axial type. (The arrangement of the vanes of the rotor and the stator is similar to that of a gas turbine.)

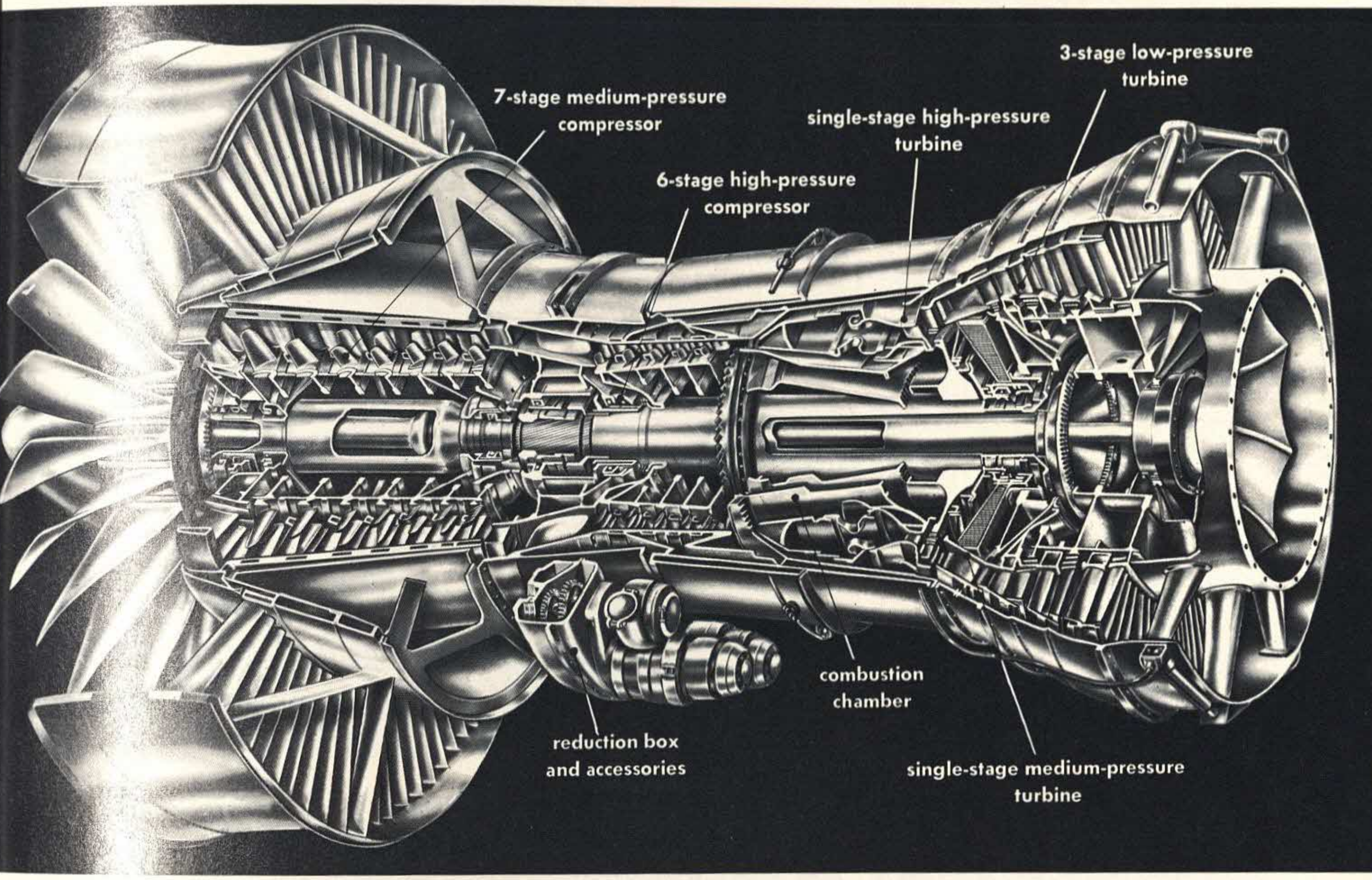
The compressed air then passes into a combustion chamber in which the fuel is burned, thus creating an increase in the temperature and the specific volume of the gas. The gas is then partially expanded through the turbine, which generally consists of a limited number of stages. The function of the turbine is to drive the compressor and other auxiliary components.

Pressure energy is converted into kinetic energy in the exhaust tube. As in the case of the rocket engine, the thrust here is again equal to the product of the time rate of exhaust gas mass expelled multiplied by the exhaust speed v; that is, $F = mv/t$.

Illustration 2b shows the theoretical thermodynamic cycle. The curve **1-2** corresponds to the compression—the compression of the first part of the curve being obtained by means of the air inlet. Line **2-3** refers to the combustion that leads to an increase of the specific volume and the temperature. This phase is followed by the expansion **3-4**, the first part of which is obtained in the turbine and completed in the exhaust tube. Point **4** corresponds to the exit of the gas into the atmosphere. The line **1-4** is a purely theoretical line. It closes the cycle and could represent the dilution of the gas expelled into the atmosphere; in fact, air at atmospheric conditions is drawn in at point **1.** For the sake of convenience in studying the operation of the engine, the cycle is here represented as closed. In actual practice, however, the cycle is open because the active fluid is always changed from one cycle to the next. Moreover, modern jet engines generally contain

THE TURBOPROP

The turboprop is similar to the turbojet but has a propeller in the front of the gas intake. This propeller is driven by the turbine through reduction gearing. The turbine drives the compressor at its own rotation rate. More of the energy in the



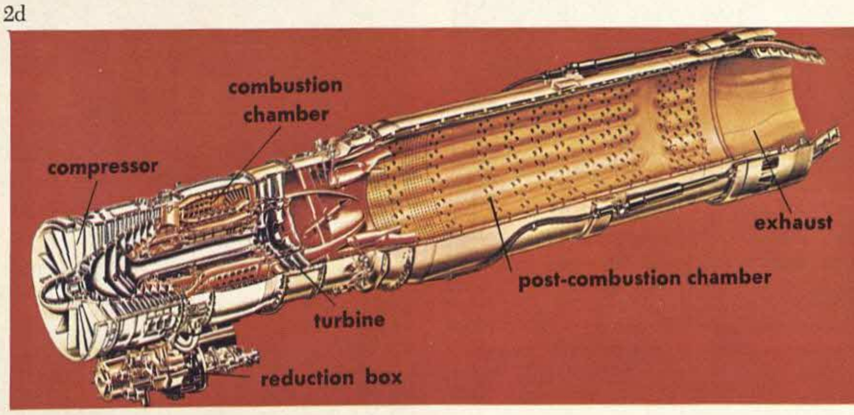
2c

modifications of the simple layout just described; the modifications are designed to augment the thrust that can be supplied by the engine.

Illustration 2c is a turbofan (bypass) engine. A part of the air drawn in by a large fan is bypassed to the exhaust tube (that is, made to expand) without first being heated. In a certain sense this creates a hybrid—something intermediate between a turbojet and a propeller engine—and obtains an increase in propulsive efficiency at low flying speeds.

Another modification of turbojet engines includes additional burners (afterburners) that are inserted between the turbine and the exhaust tube, thus obtaining an increase in the exhaust speed and, consequently, an increase of the thrust supplied by the engine. Illustration 2d shows a cutaway view of such an engine.

2d



combustion gases is used to drive the turbine than in the turbojet.

OTHER HYBRIDS

Similar in design and operation to the turbojet are other types that operate on more complex cycles and are based on technical solutions that aim at increasing thrust. These types include turbojets (for example, turboramjets) with a post-combustion cycle; that is, engines in which a second series of burners (afterburners) is inserted between the turbine and the exhaust nozzle, causing a further increase in the speed of the gases. In this case, the increase of the thrust is accompanied by a loss of efficiency. Another hybrid engine is the turborocket, whose rocket engines replace the jet combustion chamber; their exhaust drives the turbine that drives the rotary compressor. Air that is received at the intake and compressed then mixes with the rocket exhaust gases.

NUCLEAR EXPLOSION

the process and the causes of its devastating effects

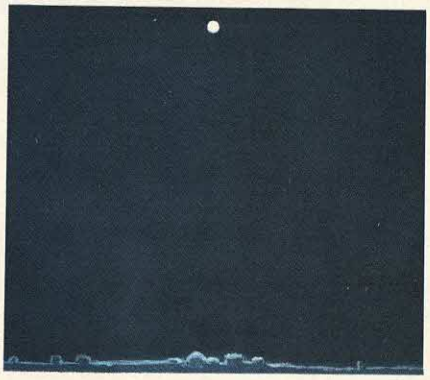
Does a nuclear explosion mean certain death to anyone caught in the vicinity of its center? The answer: something *can* be done to minimize the dangers created by such an explosion; but first, it is necessary to have a thorough knowledge of what goes on while the explosion is actually taking place. Contrary to popular belief, a nuclear explosion is not instantaneous but involves a matter of several seconds. Only by knowing the phenomena that occur during the explosion phase can meaningful protective measures be taken.

T_0—The nuclear explosive mass makes its "critical" turn—the beginning of the chain reaction that leads to disintegration—at the instant of explosion (Illustration 1).

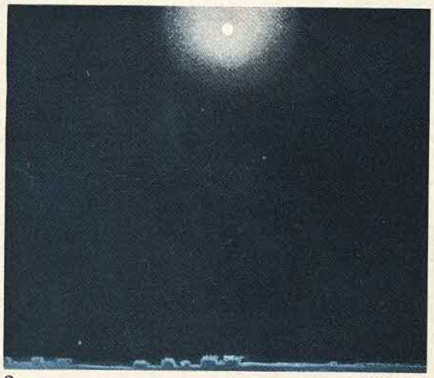
T_0 + 0.001 SECONDS—The exploding gas mass becomes enormously hot, increases in diameter, and emits an enormous quantity of x-rays and ultraviolet rays (Illustration 2).

T_0 + 2 SECONDS—The exploded mass and the surrounding air, having absorbed the energy from the explosion, form a "fireball" of intense heat and brightness (Illustration 3), which in a few seconds attains a diameter of 2.5 km (about 1.5 mi).

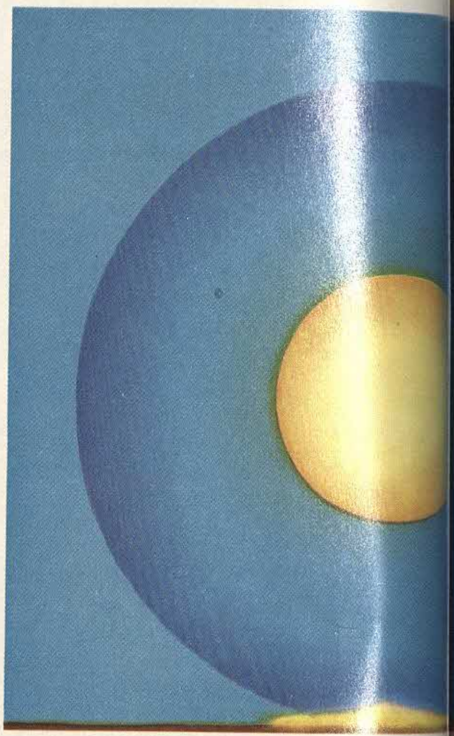
1



2



4



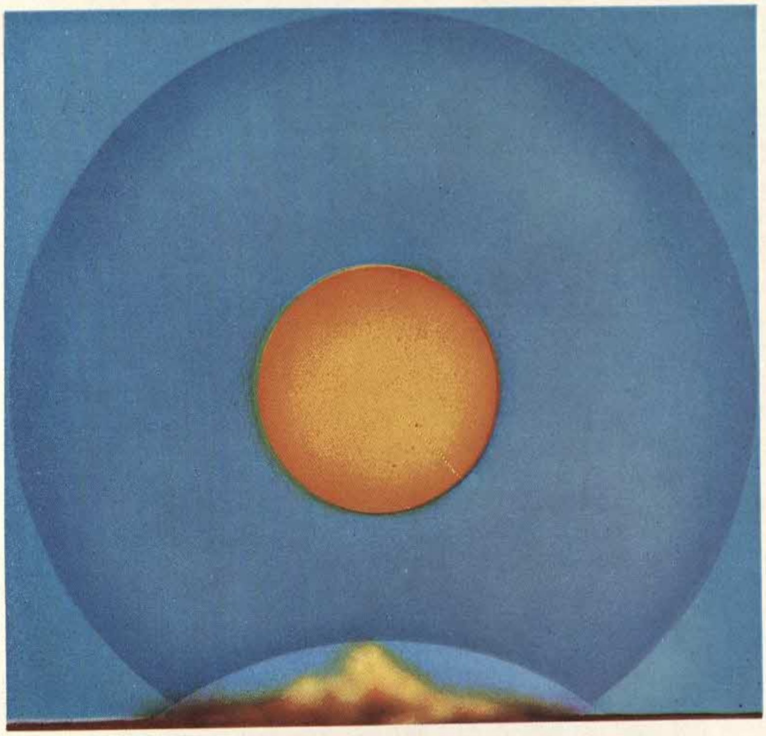
3

T_0 + 6 SECONDS—The fireball (Illustration 4) still radiates considerable amounts of energy, but its temperature has fallen appreciably, and the dangers of permanent damage to the sight of humans who see it are negligible. The shock wave reaches the Earth and produces the first mechanical damage. It is possible to protect oneself at this time by a barrier or by covering exposed parts of the body with light-colored clothing.

T_0 + 13 SECONDS—Fireball temperature and hazards continue to decrease, but damages from shock wave and the succeeding explosion blast increase (Illustration 5).

T_0 + 30 SECONDS—At this time (Illustration 6), the effects of thermal and light radiation cease, and the characteristic mushroom cloud begins to form and rise.

5



6

T_0 (THE INSTANT OF EXPLOSION)

Suppose a 1 megaton bomb, equivalent to the explosive energy of a million tons of TNT, is exploded at an altitude of 2,000 m (about 6,500 ft). At zero time T_0 (the instant of explosion), a chain reaction is triggered that leads to the disintegration of the critical mass of the nuclear explosive. This reaction occurs because the neutrons that have been liberated in the initial fission produce other fissions of nuclei situated in the same mass of the explosive. Because of this chain reaction, the explosion takes place in an extremely short time interval; the neutrons are very "fast," and in the course of about a millionth of a second the bomb mass liberates approximately 1,000 billion (10^{12}) kcal (kilocalories) of energy. (One kcal equals 1,000 calories, a calorie being a unit of energy defined as the quantity of energy or heat necessary to raise the temperature of 1 gram of water by 1 C° at standard atmospheric pressure.)

The temperature of the gaseous mass into which the bomb is converted instantly rises to several million degrees, and the gas pressure reaches a million atmospheres. One atmosphere, the standard atmospheric pressure, is the pressure of air at sea level, about 14.7 pounds per square inch, or about 1 ton per square foot. (The gas pressure of the nuclear explosion is about 147,000,000 psi.)

$T_0 + 0.001$ SECONDS

The exploded gas mass increases in diameter, emitting an extremely large quantity of x-rays and ultraviolet rays. The x-rays are absorbed into the atmosphere and discharged in the form of ultraviolet, light, and infrared rays. This intense emission of electromagnetic radiation (amounting to three-quarters of the total energy released by the explosion) causes the "initial flash" of the explosion.

The brilliance of this flash can cause total blindness to any person within a radius of several dozen miles, because the amount of light emitted per unit of surface of the luminous sphere is initially tens or hundreds of times greater than that of the surface of the sun.

Because of the atmospheric absorption of radiation, the surrounding air becomes enormously hot. After about a thousandth of a second, the luminous mass containing volatilized matter from the bomb expands into a sphere nearly 500 ft in diameter. The time elapsed since the moment of explosion is so brief that anyone in the vicinity of the explosion could have done nothing to protect himself, not even having time to close his eyes.

$T_0 + 2$ SECONDS

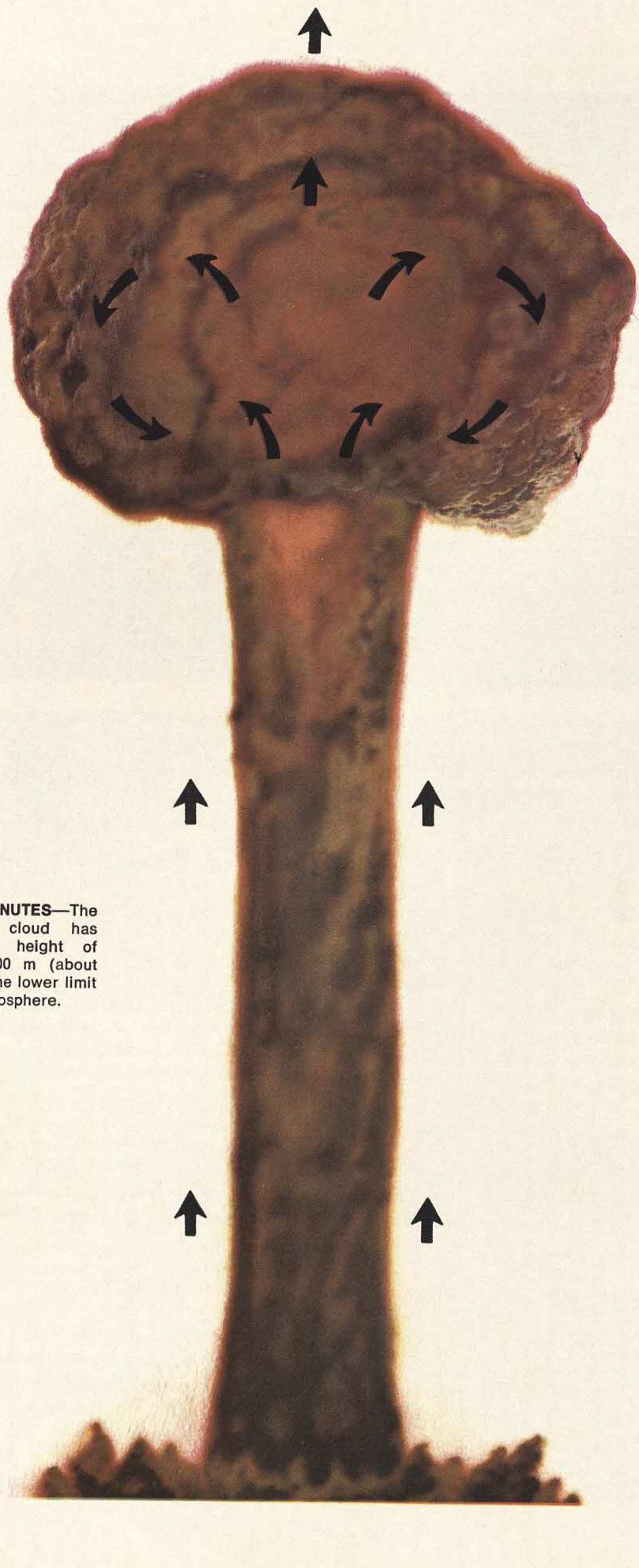
The exploded mass and surrounding air form a "fireball" whose surface, still extremely hot and as bright as or brighter than the sun's surface, produces the major effects of thermal destruction. The radiating heat is sufficient to burn all flammable materials within a radius of 4-5 km (about 2.5-3 mi). The brilliance of the fireball is still sufficient to cause permanent damage to the sight.

Around the fireball, in the meantime, a shock wave has developed, traveling at supersonic speed. This is the compression wave of the air displaced by the initial explosion.

$T_0 + 6$ SECONDS

The shock wave then strikes the ground, where it causes the first mechanical damage. The wave causes a violent compression of the air, the intensity of which decreases as the distance from point zero (the explosion center) increases. At about a mile from point zero, the added

7



T_0 **+ 2 MINUTES**—The mushroom cloud has reached a height of about 12,000 m (about 40,000 ft), the lower limit of the stratosphere.

pressure may be about 2 tons per square ft ($2\,t/ft^2$)—about double standard atmospheric pressure. A human has an even chance of surviving exposure to pressures of less than $4\ t/ft^2$. Chance of survival is about 1 in 100 when pressure exceeds that amount.

In the early seconds after the instant of explosion, it is vitally important that the body be protected against heat radiation. Even in a period of this short duration it is possible to find shelter behind a barrier or, at least, cover those parts of the body not protected by clothes. Clothing, especially if of light color, provides some protection against the heat radiated by the fireball. Although clothing may catch fire, injuries so caused will be less serious than injuries that could be suffered if the burning material is removed and the body exposed to direct radiations.

$T_0 + 13$ SECONDS

The shock wave spreads along the ground surface and is followed by the so-called explosion "blast" caused by displacement of the air expelled by the fireball. This blast sweeps along the ground at a speed of between 300 and 400 km/hr (about 180-250 mph). After the sudden increase in pressure from the shock wave, pressure drops to about one atmosphere. Actually, the removal of the air from the center of the explosion gives rise to a vacuum.

Meantime, the fireball has cooled and diminished in volume. Being lighter than air, it rises in accordance with Archimedes' principle of buoyancy. This upward movement produces a reversal of wind direction on the ground; after having first blown outward from the center of the explosion, the blast now comes rushing back toward the center.

The shock front pressure causes heavy damage to closed structures (such as buildings with few or small window openings), while leaving relatively intact such open structures as electrical transmission towers, cantilever bridges, and glass-and-steel skyscrapers. All such structures undergo internal damage from ground blast and the blast reversal. Vast quantities of powdery dust are raised in the vicinity of the blast.

$T_0 + 30$ SECONDS

As it rises, the fireball loses its characteristic spherical shape and takes on the typical mushroom shape. The arrows in Illustration 7 show that the mushroom shape evolves because of turbulence inside the fireball.

$T_0 + 2$ MINUTES

The mushroom cloud has now reached a height of 12,000 m (about 40,000 ft), about the lower limit of the stratosphere.

Winds blowing at such high altitudes gradually dissipate the mushroom cloud, and the materials contained in it, commonly known as fallout, are dispersed into the atmosphere. Since these materials are now very fine particles held in suspension by hot gases, they can rise to still greater heights and be carried by the winds around the Earth several times before they settle.

MECHANICAL EFFECTS

When a nuclear explosion of the order of 1 megaton takes place, the mechanical effects of the explosion (produced by the pressures of the shock wave and explosion blast) make up about 50 percent of the destructive effects. Within a radius of about 3 km (about 2 mi) from ground zero (the ground point beneath point zero), almost total destruction of all buildings, bridges, highways, and railroad installations occurs. At about 4.5 km (about 2.8 mi) from ground zero, 50 percent of humans suffer bursting of both eardrums. At about 12 km (about 7.5 mi), the damages to buildings are not of serious importance. At 25 km (about 15 mi), buildings show signs of damage to doors and windows. Only at 50 km (about 30 mi) from ground zero are damages negligible.

THERMAL RADIATION EFFECTS

Thermal radiation is responsible for about 35 percent of all the damage produced by a nuclear explosion. At a distance of less than 4.5 km (about 2.8 mi) from ground zero, any flammable material catches fire. At distances of 13, 16, and 22 km (about 8, 10, and 14 mi), the human body suffers third, second, and first degree burns, respectively, when exposed to radiations emanating from the fireball.

RADIOACTIVE FALLOUT DAMAGE

At a comparatively short distance from ground zero—about 1 km (about 0.6 mi) —radioactive effects of nuclear radiations are felt.

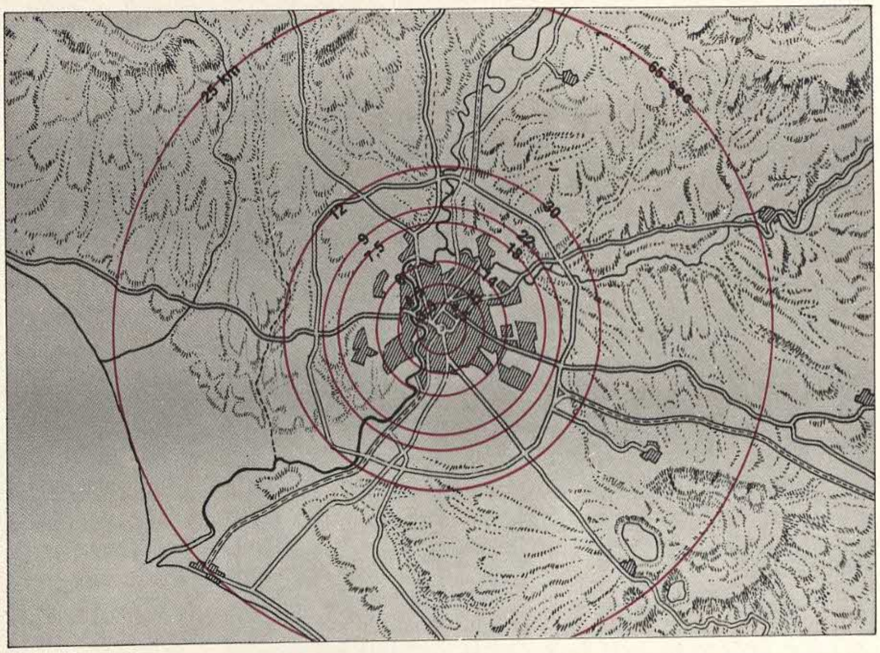
Although at greater distances these radiations are absorbed by the atmosphere, the radiations emanating from dust particles of "activated" materials (substances made radioactive by the explosion radiations) are still dangerous.

Direct radiations within a radius of about 1 km from ground zero are lethal. At twice that distance the death toll is about 50 percent, while at three times the distance the immediate effects of such radiations are negligible.

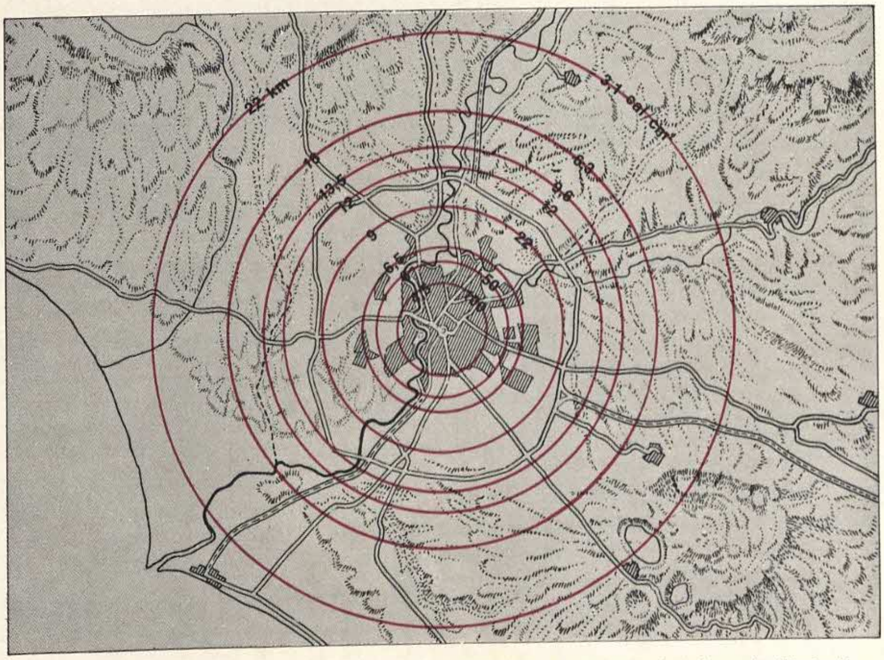
The radioactive dust, or fallout, falls partially within a brief period of time at a short distance from ground zero. Part of this dust remains in suspension in the stratosphere, from which it drops back to Earth in rain or other precipitations. This return may take a number of years. In the absence of particularly strong winds, particles of more than 10 μm (microns) in diameter usually fall back within 24 hours. The manner in which the particles spread over the ground cannot be accurately predicted.

Generally, the greater part of the fallout occurs in the direction in which the wind blows the dust. Knowing this fact may be useful in determining which regions should be avoided in the period immediately following a nuclear explosion.

MECHANICAL EFFECTS—If a nuclear charge of 1 megaton (10^6 electron volts) exploded over a city (Illustration 8a), the mechanical effects of the explosion would be appreciable up to a maximum distance of slightly more than 50 km (about 30 mi).



8a



8b

THERMAL EFFECTS—The thermal effects of the same explosion would be felt within a smaller radius than that for the mechanical effects, but the thermal effects are just as destructive. The explosion shown in Illustration 8b occurred at approximately 2,000 meters (about 6,700 ft) above the ground.

NUCLEAR FUELS | fuel elements used in heterogeneous reactors

Nuclear reactors are classified as homogeneous or heterogeneous according to the way the fissionable material (fuel) and the moderator or coolant are distributed inside them. In a homogeneous reactor, fuel and moderator are finely divided and evenly mixed together. In a heterogeneous reactor, the two substances are isolated; the fuel is in the form of rods or plates arranged in a lattice with the moderator occupying the space between. Most nuclear reactors, including the first successful one assembled in Chicago in 1942 and the industrial-scale nuclear power plants now in use, are of the heterogeneous type.

The rods or plates of fuel that go into the heterogeneous reactor are the fuel elements. When the nuclear reactor is in operation, heat and fission products are generated within these elements. A liquid or gaseous coolant must circulate, to remove the heat quickly and safely from the reactor core before dangerously high temperatures build up. To protect the fuel from the chemical action of the coolant and to keep fission products from polluting the coolant, the fuel elements are clad with a thin sheath or coating of some heat- and corrosion-resistant metal. The fission products generated within the fuel element are largely gaseous; they include krypton, iodine, and xenon. The sheathing must be able to withstand the pressure that builds up inside the fuel element as these gases accumulate. In thermal (slow) neutron installations, including the industrial-scale power plants, the fission products are retained within the fuel elements. In fast reactors, a special draining system is needed to remove the fission products; it uses a current of inert gas that passes across the fuel elements and extracts the fission gases from them.

The fuels that have been used in nuclear reactors to date are chiefly metallic uranium and uranium oxide (UO_2). The former has been widely used in gas-graphite reactors, while the latter has been used primarily in water-cooled reactors. The dioxide is only slightly reactive with water, so that no serious problems develop if a break in the sheathing allows the fuel to contact the water.

Studies are being carried out on other substances that can be used as nuclear fuels. These include uranium carbide (UC) and uranium silicide (U_3Si). The former is adaptable to cooling with organic fluids, and can be used in high temperature systems. The silicide may prove useful in water-cooled systems because of its low level of reaction to water.

Important characteristics of the fissionable materials used as nuclear fuels are their densities and maximum working temperatures. High density yields high power per unit of volume and makes possible the design of relatively small, compact installations. High working temperature makes possible an efficient thermodynamic cycle.

TABLE 1

PROPERTIES OF SOME FISSIONABLE SUBSTANCES

Substance	Density, g/cm³	Maximum safe temperature, °C	Melting point, °C
metallic-U	19.0	668	1,130
UO_2	9.7	2,800	2,800
UC	13.6	2,000	2,350
U_3Si	15.6	930	1,650

Table 1 gives the densities, maximum working temperatures, and melting points for the four fuels mentioned. Maximum working temperatures of a reactor must be less than the melting point of the fuel to prevent an unacceptable phase transformation.

There are also plutonium fuels containing plutonium extracted during the regeneration of a primary fuel (^{238}U is changed into plutonium through neutron radiation and the subsequent radioactive decay). The plutonium may be used alone or in combination with uranium; fuel elements made of a mixture of uranium and plutonium oxides have been given special attention. Mention should also be made of thorium, a fertile element from which a fissionable fuel can be obtained (thorium-232 can capture a neutron to form thorium-233, which decays into uranium-233).

The material used to clad or sheath the fuel element must resist corrosion and high temperatures; it must also be a material through which neutrons can easily pass. Table 2 lists some of the materials used for this purpose.

Beryllium has not been widely used because of its low ductility. Magnesium is used only in alloys with aluminum for the shields of the fuel elements in gas-cooled, metallic-U reactors. Stainless steel has been widely used as a sheath for fuel elements, although it is not listed in Table 2. Stainless steel is quite strong and resistant to corrosion, but its neutron absorption section is so high that it is generally used only with fuels that have been enriched with fissionable isotopes.

TABLE 2

PROPERTIES OF SOME SHEATHING MATERIALS

Material	Absorption section, barns	Density, g/cm³	Melting point, °C
beryllium	0.009	1.85	1,278
graphite	0.0037	1.9	3,845
magnesium	0.069	1.74	651
aluminum	0.24	2.7	660
zirconium	0.18	6.5	1,852

1

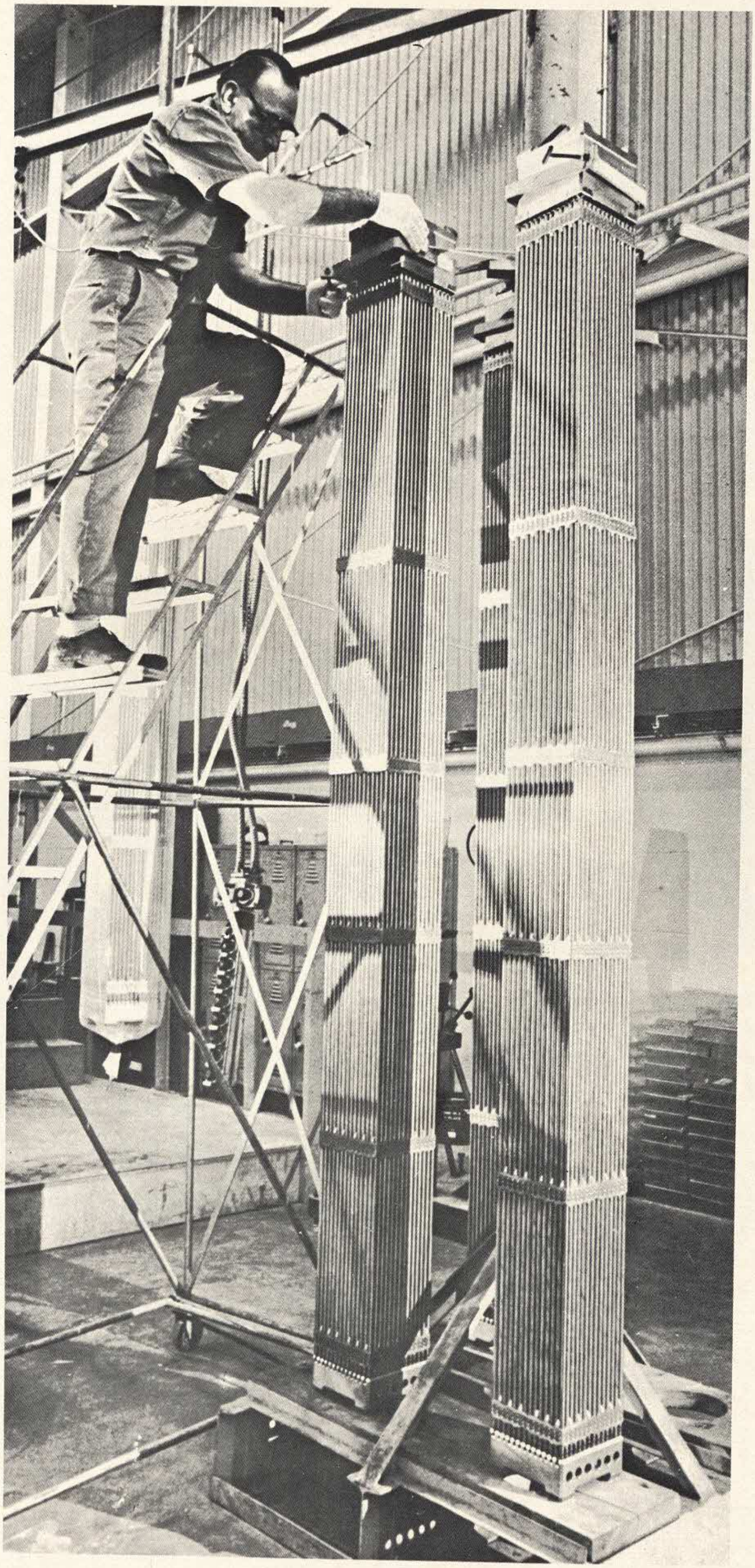
GAS REACTORS—The fuel elements for gas-graphite reactors of the ordinary or Calder Hall type (named for the first British nuclear power plant) consist of rods of natural uranium. Each rod is about 1 m (about 3 ft) long and 2.75 cm (about 1 in.) in diameter. Illustration 1a shows a device that checks the dimensions of the uranium rods. The rods are sheathed with a magnesium-aluminum alloy. Its low melting point and that of similar alloys place a limit on the temperature of the gas coolant, a maximum temperature of about 450° C (842° F). Illustration 1b shows the fins that are fitted around the sheathed rods to improve heat transmission.

The fuel elements for gas reactors that work at higher temperatures, such as the AGR (advanced gas-cooled reactor) working at 650° C (1,202° F), consist of sintered uranium oxide rods sheathed in stainless steel and grouped in bundles.

At still higher temperatures, as in reactors where helium emerges from the core at a temperature of 750 to 800° C (1,382 to 1,472° F), the fuel elements are sheathed with impermeable graphite. Illustration 1c shows the assembling of a 3.6 m (about 12 ft) fuel element for the HTGR (high temperature gas reactor) at Peach Bottom, Pennsylvania. The ring-shaped units are made of sintered, enriched uranium oxide; they are threaded on a graphite rod and inserted within the tubular graphite sheath. A graphite stopper, which is used to seal the end of the fuel element, is shown to the right of the long tube.

Illustration 1d shows a device developed to handle the completed fuel elements from above.

2a



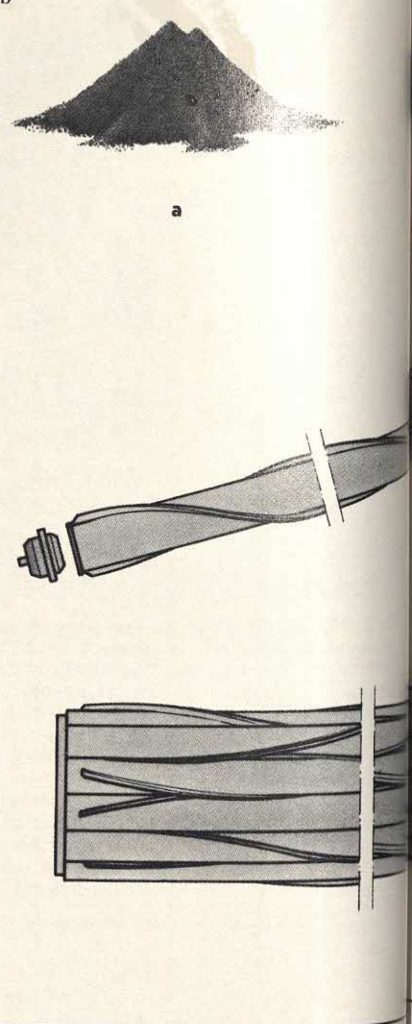
WATER REACTORS—In light water reactors of the PWR type and the BWR type, the fuel elements consist of bundles of slender rods, each as long as the core of the reactor. The rods are assembled into a unit that has a square cross section. Illustration 2a shows a fuel element containing 204 rods. The rods are usually 1.1 to 1.2 cm (about 0.5 in.) in diameter; the length varies from 2 to 3 m (6.5 to 10 ft) according to the size of the reactor in which they are used.

The rods making up the fuel element of the water reactor are hollow tubes made of stainless steel or of a zirconium alloy, which usually contains 1.5 percent tin and 0.3 percent iron, chrome, and nickel or a similar alloy that contains no tin. The tubes are filled with pellets of sintered uranium oxide. The uranium is enriched in uranium-235 at concentrations that vary from 2 percent to 4 percent.

The natural uranium fuel elements used in heavy water reactors are also arranged in bundles, but because they are rods of larger diameter, each bundle contains fewer rods.

The small diameter of the enriched uranium rods used in the light water reactor is pos-

2b

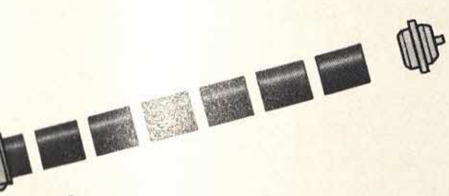


sible because of their greater power density, which derives from the presence of more uranium-235 atoms in the fuel. This means that, given a certain surface temperature for the shield, the working temperature limit in the center of the rod is attained with a smaller diameter.

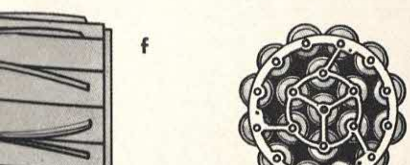
Quite often the bundle of rods has a length that is only a fraction of the height of the core, so that the power channel consists of a series of bundles placed end-to-end. This makes it possible to vary the positions of the bundles during loading operations, thus considerably improving overall utilization of the fuel by making possible a more uniform radiation during the fuel element's life in the reactor.

Illustration 2 is a schematic representation of the manufacture of the fuel elements used in heavy water CANDU (Canada Deuterium Uranium) reactors. The uranium oxide powder **a** is pressed **b** and sintered **c** and **d** into pellets. After grinding, these pellets are loaded into zirconium alloy tubes **e**, which are then sealed at one end and grouped into bundles **f.**

b **c** **d**



e

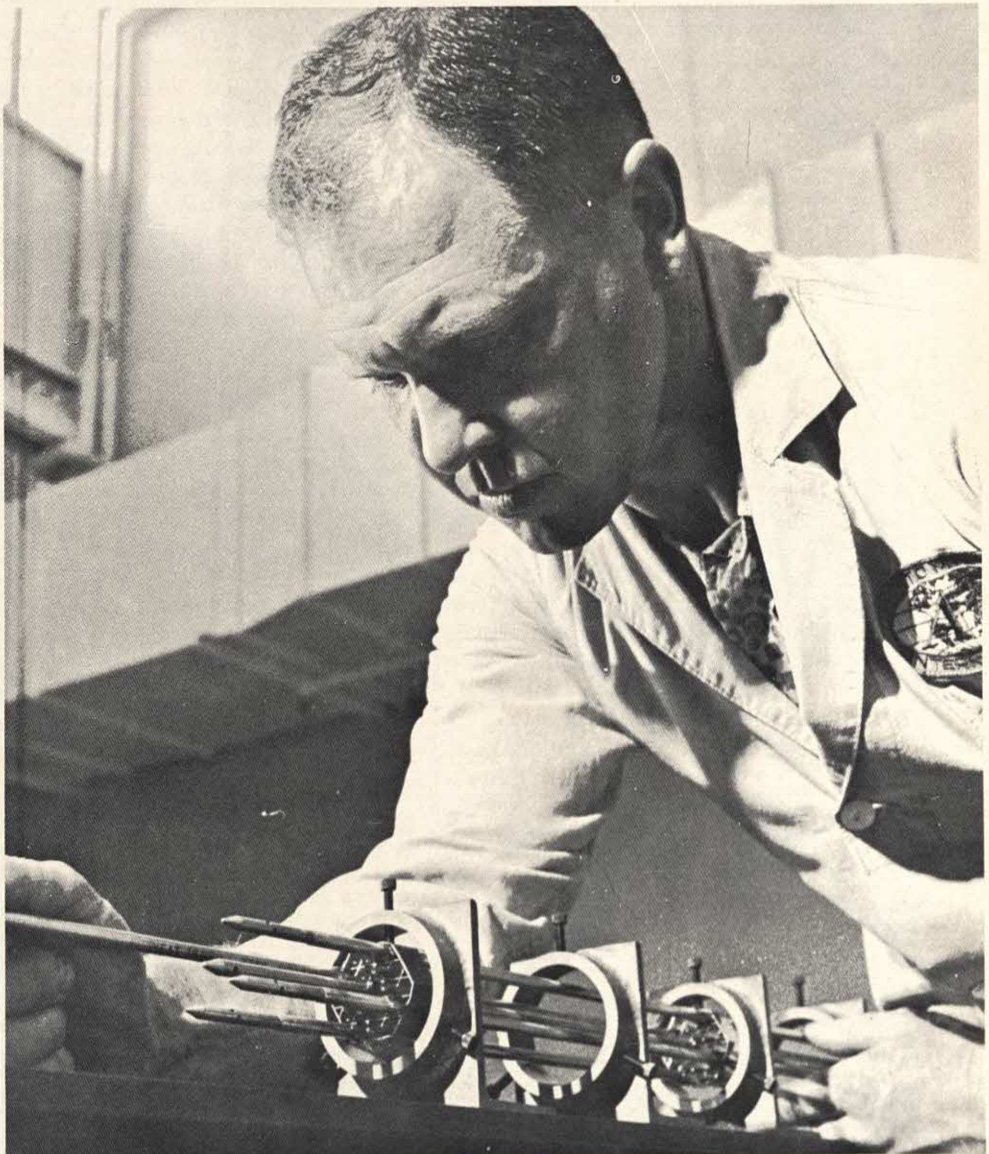


f

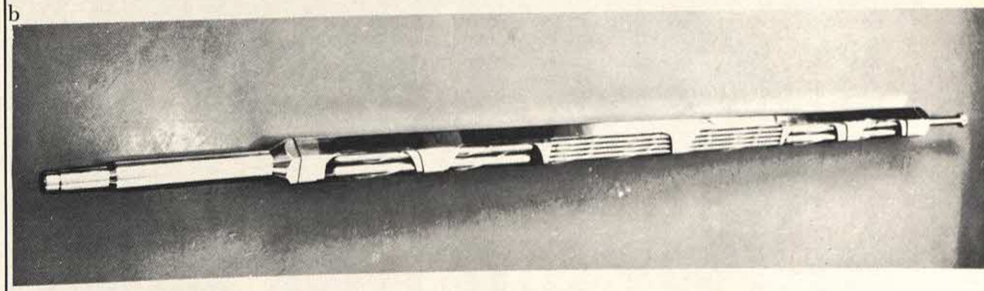
3

FAST REACTORS—In fast reactors as well as thermal ones, the fuel elements are usually bundles of rods. Fast reactors are characterized by a very high power density; their fuel is strongly enriched (20 percent or more) with uranium-235 or plutonium. As a result, the fuel rods are very slender, normally less than 10 mm (about 0.4 in.) in diameter. A typical fuel rod is 6 mm (about 0.2 in.) in diameter. Illustration 3a shows the assembly of a bundle of rods for use in the fuel element of a fast reactor. Illustration 3b shows a fuel element for the fast reactor.

a



b



THE NUCLEAR REACTOR | the giant that harnesses the energy of the atomic nucleus

The nuclear reactor is a device within which a nuclear chain reaction is initiated and controlled. In the nuclear reactors now functioning, that reaction is nuclear fission, the splitting of a heavy atomic nucleus to create two or more nuclei that are less heavy. The by-products of this reaction are heat energy and a rapid flow of neutrons. The heat energy can be converted into electricity; the neutrons can be used in the production of isotopes and in other research activities. There are, however, technical reasons why it is impossible to build a nuclear reactor that serves both purposes equally well. Today's nuclear reactors are specifically adapted to one purpose or the other; they may be distinguished as power reactors, which produce energy, and research reactors, which produce neutrons for scientific uses.

Nuclear reactors may be constructed to use a number of different fissionable materials or fuels; among them are uranium, thorium, and plutonium. They may also use a number of different moderators; that is, materials capable of slowing down the neutrons. A still greater variety of materials, including base metals and alloys, is used as structural material in the reactor itself. The design of the reactor determines the way in which fission takes place within it.

THE FISSION OF THE NUCLEUS

The fission reaction takes place when an extra neutron is captured and absorbed by the nucleus of an atom, causing the nucleus to split into two or more fragments while releasing energy and more neutrons that may, in turn, be absorbed by other nuclei and trigger additional reactions. The uranium-235 nucleus is one in which this reaction may be caused by the absorption of either a slow-moving neutron (one with low kinetic energy) or a fast-moving neutron. The uranium-235 nuclei do not always split in the same way (producing identical fragments and the same number of neutrons); but the new nuclei are always lighter than the uranium-235 nuclei, and the number of neutrons released is always greater than one. It is generally true that:

$$^{235}U + 1 \text{ neutron} \rightarrow \text{fission fragments} + n \text{ neutrons} + \text{energy} + \text{radiation}.$$

The significant fact is that the number n is always greater than 1; it is, on the average, equal to 2.5. Only one neutron is needed to trigger the fission reaction. Once the reaction has taken place, however, more than one neutron—typically, two or three neutrons—are free and capable of triggering new reactions if they happen to be absorbed by the nuclei of uranium atoms. In this way, the chain reaction is initiated.

THE UNCONTROLLED CHAIN REACTION

The neutrons emitted when the fission reaction takes place possess tremendous energy and move at extremely high velocity. As they move away from their points of origin, they strike a large number of nuclei, lose part of their energy, and are gradually slowed. As they move slowly, the possibility of their being captured and absorbed by a nucleus increases; when moving at high speed they may not be in contact with a nucleus long enough for the reaction to occur.

In a mass of uranium-235 of infinite dimensions one fission reaction would inevitably initiate a chain reaction. This is because all of the neutrons emitted by the first reaction would be certain to remain within the mass until they slowed and were captured and absorbed by other nuclei; each neutron would then cause a new fission reaction. The chain reaction would snowball, and the entire mass would disintegrate.

In a very small mass of uranium-235 the chain reaction might not be initiated at all because all of the neutrons emitted by the first fission reaction might escape from the mass before being captured and absorbed by other nuclei.

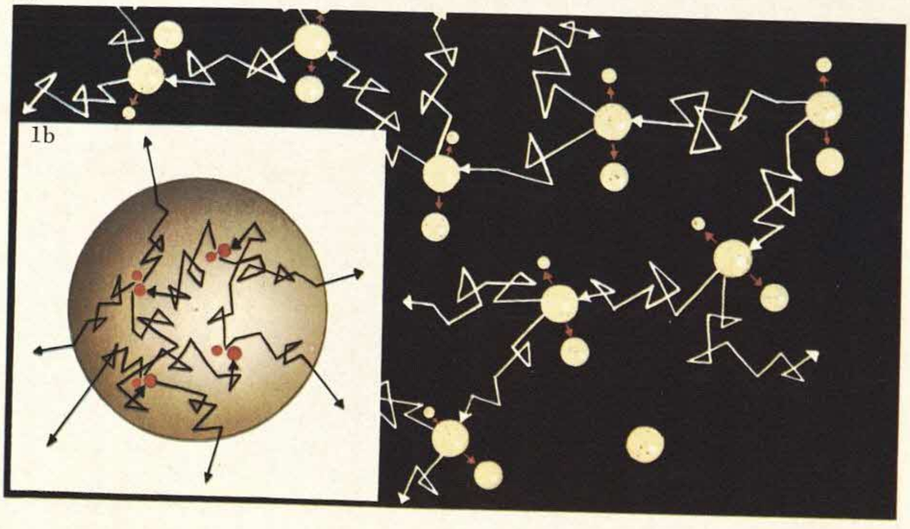
Between these two extremes there is a value for the mass of uranium-235 that is the minimum in which one fission is certain to trigger a chain reaction by causing more fission. This is the critical mass. For uranium-235 the critical mass is slightly more than 7 kg (about 15 lbs). The critical mass varies from one fissionable material to another. For plutonium, it is slightly more than 1 kg (about 2.2 lbs).

In a mass of fissionable material equal to the critical mass or larger, the chain reaction may go out of control and explode.

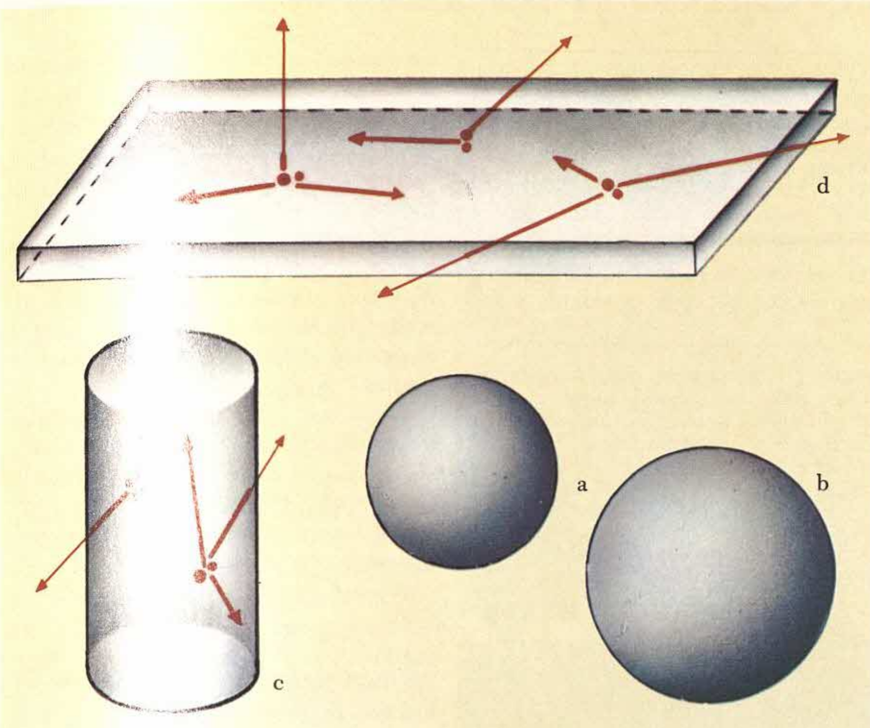
THE CONCEPT OF CRITICAL MASS—In Illustration 1a, a fission chain reaction takes place within an infinitely large mass of uranium-235. The large spheres represent uranium-235 nuclei; the smaller spheres represent the unequal fragments into which they split. The arrows represent the neutrons emitted. Every neutron remains within the mass until it is absorbed by a nucleus, and the chain reaction continues until all of the uranium has disintegrated.

Illustration 1b shows the same process within a sphere of limited dimensions. Some of the neutrons escape. If too few remain, the reaction dies out. The smallest sphere in which the reaction will be self-sustaining is the critical mass.

1a



2



CRITICAL MASS IN DIFFERENT GEOMETRICAL SHAPES—A mass of fissionable material of the dimensions shown in Illustration 2a is not critical. The one shown in Illustration 2b is of critical mass. The cylindrical mass in Illustration 2c is larger than in 2b, but it is not critical because it has more surface area that enables many neutrons to escape from it. The mass of the flat sheet in Illustration 2d is even larger, but its surface area is so great that nearly all the neutrons escape from it.

THE CRITICAL STATE

The value of the critical mass depends on a number of factors. The figures given for uranium-235 and plutonium are valid only when the substances are pure and when they are arranged at maximum density in a sphere.

The sphere gives maximum volume with minimum surface area. In this case the volume represents space within which the neutrons may unite with nuclei and trigger new reactions; the surface represents a limit beyond which the neutrons are lost as far as fission reactions are concerned. The sphere, therefore, represents the minimum critical mass of an isolated fissionable material.

Suppose 7.5 kg (about 16.5 lbs) of uranium-235 were shaped into a thin, flat sheet rather than into a sphere. A sheet has much more surface area than a sphere, and the neutrons emitted by the first fission reaction would be much more likely to emerge from the mass before being captured by nuclei and triggering new reactions. Obviously, the mass of the flat sheet will have to be much greater than the mass of the sphere if the critical state is to be achieved. In general, the value of the critical mass depends on shape.

Suppose a sphere of fissionable material were surrounded by some material that reflected back into the mass the neutrons that escaped from it. In this case, the mass required to achieve the critical state would be less than if the sphere of fissionable material were surrounded by a vacuum or some nonreflective material. Here, the concept of critical mass has been enlarged to that of critical system: the system consisting of the mass of fissionable material and the reflector surrounding it. Slightly more complex systems are found in nuclear reactors, where the fissionable material may be combined with or surrounded by a moderator intended to slow down the neutrons and increase the likelihood of their uniting with nuclei.

The system within a nuclear reactor may be described in the following ways:

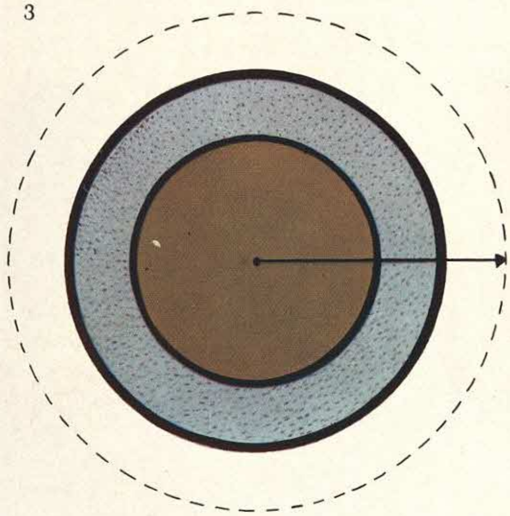
1. the system is subcritical if it cannot support a chain reaction;
2. the system is critical if it will support a chain reaction that can be controlled so it will neither die away nor increase spontaneously;
3. the system is supercritical if it will allow a chain reaction to take place at increasing speed so that it goes out of control and becomes an explosion.

THE NUCLEAR REACTOR

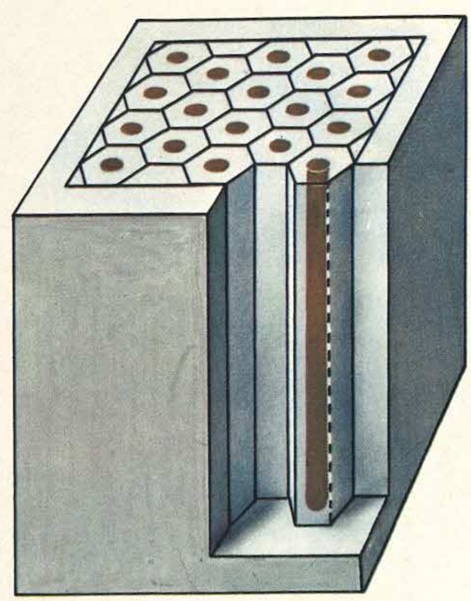
A nuclear reactor contains fissionable material such as uranium or plutonium combined with a moderator such as graphite or heavy water that can slow down the neutrons. The two materials may be mixed together, or they may be in separate rods or cylinders. Some of the neutrons generated by fission may be absorbed by the moderator, and some may escape through the casing of concrete or some other dense material that

THE REFLECTOR—A quantity of fissionable material that is less than the critical mass may become critical if it is surrounded by some material such as graphite that turns back the neutrons that would otherwise escape from it. The dotted line corresponds to the dimensions the sphere would need to be critical without the reflector.

3



4



FUEL RODS—The core of a nuclear reactor may consist of a number of rods of fissionable material (fuel) surrounded by a moderator such as graphite that slows down the neutrons so that they are more likely to produce more fission reactions.

A PART OF THE FIRST NUCLEAR REACTOR —The first successful nuclear chain reaction was achieved by the Italian-American physicist Enrico Fermi on December 2, 1942, in Chicago. Some blocks from that original "pile" were used in constructing this subcritical reactor at the Argonne National Laboratory, which is near the city of Chicago. The reactor is used in the training of nuclear physicists, chemists, and engineers.

5

surrounds the reactor. If a chain reaction is to be sustained, at least one of the approximately 2.5 neutrons generated by each fission reaction must trigger a new reaction. This condition may be expressed as follows:

The number of neutrons produced within the reactor in a given instant $N \times$ the number of neutrons K not absorbed by the moderator $\times$ the number of neutrons P that do not escape from the reactor = 1, or: $N \times K \times P = 1$.

When this condition exists, the chain reaction continues but without increasing. The number of free neutrons circulating within the reactor, and the number of fission reactions taking place each instant, remain constant.

When the following condition exists:

$$N \times K \times P < 1$$

the number of neutrons circulating within the reactor is decreasing, and fewer reactions take place each instant. As a result, the chain reaction dies.

If the following condition exists:

$$N \times K \times P > 1$$

the number of neutrons within the reactor is increasing, more reactions are taking place each instant, and the chain reaction may go out of control.

ENERGY PRODUCED BY THE FISSION REACTION

The fission of one uranium nucleus frees energy of slightly less than 200 Mev (million electron volts). By comparison, the most vigorous chemical reactions known free less than 10 ev of energy per molecule of reaction product. The fission of one pound of uranium yields heat energy equal to that produced by burning 3 million lbs of coal. Obviously, fissionable materials are fuels capable of yielding quantities of energy vastly greater than those obtainable from conventional fuels.

Although the ores of uranium and thorium are being used rapidly, there is an essentially limitless supply of fission reactor fuel available. The past fears about exhaustion of energy supplies can be allayed, because the energy content in the rocks and the seas is believed sufficient to last for 10 billion years.

NUCLEAR TECHNOLOGY—I | particle accelerators

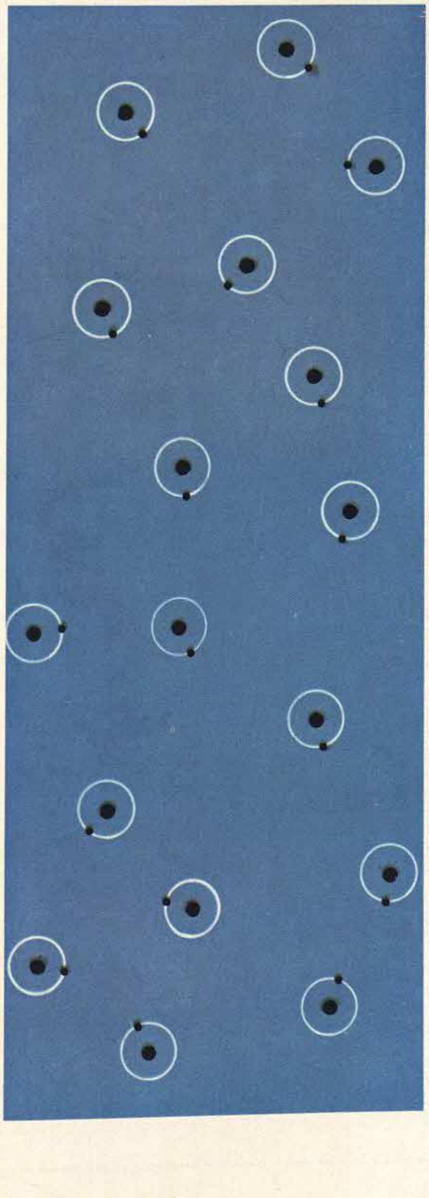
The early studies of radioactivity led to the rapid development of nuclear science. On the basis of the scattering experiments performed by the British physicist Ernest Rutherford between 1908 and 1913, radioactive transformations could be attributed to spontaneous changes of an atomic nucleus. Rutherford and his associates had discovered how the principal components of atoms reacted under bombardments of charged particles. These and following experiments led to the founding of nuclear physics within a relatively short time. As knowledge grew, physicists wanted to perform much more complicated experiments. They felt a need to know, for example, the precise composition of a beam of charged particles where the beam was hitting its target, and the dispersion of energy at the target. To obtain this knowledge, extreme control accuracies were needed. In the 1930s, the first machines to achieve this control were built—accelerators. An accelerator increases the velocity and energy of charged elementary (or fundamental) particles—for example, electrons or protons—through applications of electric and/or magnetic forces. Accelerators have made particles move at velocities approaching the speed of light. This article examines the atom and its constituents and the newly developing techniques of particle acceleration.

ATOMS AND PARTICLES

Illustration 1 is a greatly enlarged view of a small volume of hydrogen gas. The hydrogen atoms are placed far apart in order to make their structure stand out clearly. The hydrogen atom is the simplest atom found in nature. It consists of a nucleus (containing a single proton—a heavy elementary particle with a positive electric charge) and an electron that revolves around this nucleus. A proton is over 1,800 times as heavy as an electron, which carries a negative charge. The value of the electron's negative charge is equal to that of the proton's positive charge. Therefore, when these two particles are bound together in the atom, the two charges cancel out each other and the atom, as a whole, is neutral. The proton and the electron are bound together by the force of electrostatic attraction. They do not come completely together because the electron orbits around the proton and the centrifugal force due to this motion compensates for the force of electrostatic attraction (the gravitational force between the mass of the proton and that of the electron is negligible when compared to the electrostatic attraction). If the hydrogen atom collides with a suitably accelerated particle (another atom, an electron, or a proton), the electron is detached and hurled far from the proton, overcoming the force of electrostatic attraction. This process is known as ionization, and the hydrogen atom becomes an ion. It is a rather special ion because it consists simply of the nucleus of the hydrogen atom. This nucleus is the simplest of all atomic nuclei. It is an elementary particle—that is, one of the fundamental particles that constitute all the matter in the universe.

SPLITTING ATOMS IN ORDER TO STUDY THEM—If nuclear particles are accelerated to high speeds and are made to collide with the atoms of a certain substance (hydrogen, for example), the atoms split into two parts. The fragments obtained in this manner can be recognized and studied. In this particular case (a hydrogen atom) an electron and a proton are obtained. Splitting atoms, therefore, helps scientists to determine their exact structures.

1



When an atomic physicist studies the constituents of an atom, he must supply energy to it. A simple way of supplying energy to an atom is to heat the substance of which it is a part. Atoms in substances are never at rest, and they move with increasing speeds as the temperature increases. A rise in temperature, therefore, causes the atoms to move more violently and to collide with one another. Suppose, for example, that copper is heated to a temperature of about 4,000° C (about 7,200° F). This is an everyday phenomenon that occurs whenever an electric plug is pulled from its socket. Pulling the plug from the socket creates a spark in which a large number of copper atoms (in the plug) become volatilized. The copper then becomes luminous, producing a green light. This fact permits some conclusions to be drawn concerning the structure of the copper atom—for example, how some of the electrons are arranged in the various orbits around the nucleus. The same method may be used to determine the structures of atomic nuclei and elementary particles. The speed needed for studying atomic structures is much lower than that needed for studying particles. Particles are made to collide with each other or with nuclei at extraordinary speeds. Collisions must occur, not at speeds of a few miles per second (as is the case when studying atomic structures), but at tens of thousands of miles per second—even speeds close to the speed of light.

CHARGED PARTICLES

Particles can be accelerated to high speeds by exploiting one of their fundamental properties—their electric charge. If the particles are placed in an electric field, they are accelerated by this field until they attain a high speed, a speed

that depends on the following factors:

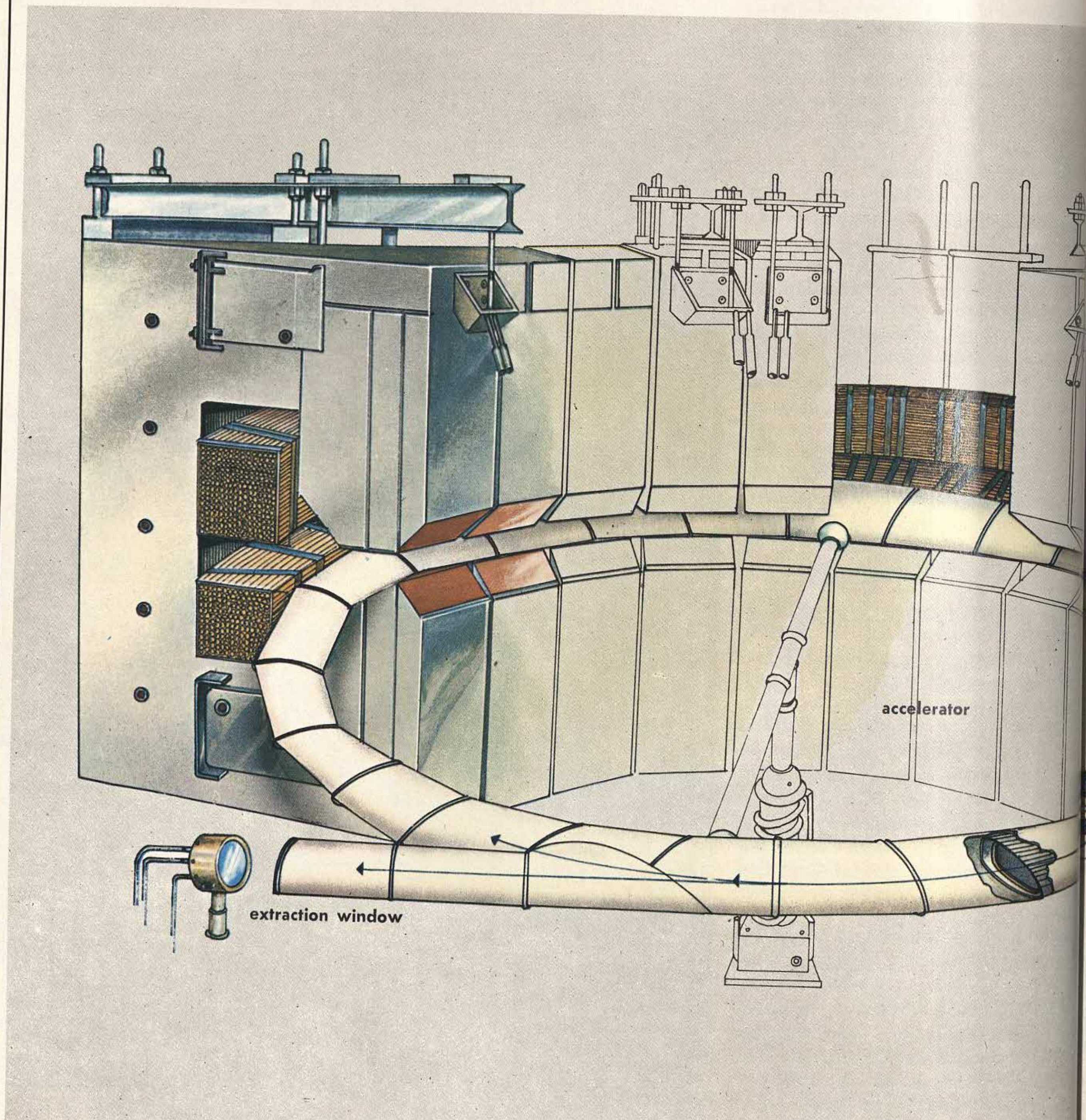
1. The intensity of the electric field: the greater the voltage the particle must pass through, the greater the speed it attains.
2. The amount of charge: a helium nucleus (alpha particle, charge 2+) becomes subject to a force that is twice as great as the force that the same field exerts on a hydrogen nucleus (proton, charge 1+).
3. The mass of the particle: the greater the mass of the particle, the greater is its inertia and, consequently, the smaller is the acceleration that a given electric field produces.

The principal particles studied by means of acceleration are protons, electrons, alpha particles, and the nuclei of elements that are heavier than hydrogen and helium (lithium, beryllium, and bo-

THE PRINCIPLE OF AN ACCELERATOR — This illustration depicts the basic structure of a particular kind of accelerator, a cyclotron. An understanding of this simplified diagram must precede any study of the construction details of the various types of accelerators.

At the outset mention should be made of the reservoir or supply of the element whose nuclei are accelerated. This supply source may

2



ron, to cite some common examples).

The particles to be accelerated can be obtained from gaseous substances (hydrogen and helium, for example) or from solid elements from which the atoms are first detached and then ionized. Finally, heating a metal to a suitable temperature can cause electrons to be emitted.

THE ENERGY OF PARTICLES

The energy that the particles acquire as they pass through the accelerator depends on the electric field to which they have been subjected. The unit of measurement used in describing the energy acquired by the particles is the electron volt. It corresponds to the energy acquired by an electron when it is displaced between two points that are separated by a potential difference of one volt. If two metallic plates located opposite each other in a container from which the air has been evacuated are connected to a battery that produces a potential of one volt, and if an electron is at rest close to the negative plate, the electron is promptly repelled by the negative plate and attracted toward the positive plate at the other side of the container. After having crossed the space between the two plates, the electron has an energy of one electron volt. Any other particle having the same charge as the electron also accelerates until it possesses the same energy as the electron—that is, one electron volt. In a given electric field, however, a proton reaches slower speeds than an electron, because of its greater mass; nevertheless, its energy is the same.

Particle accelerators are used to accelerate protons and electrons from minimum energies of a few hundred thousand electron volts up to energies on the order of tens of billions of electron volts. The speed of particles possessing energies of this magnitude must be calculated on the basis of the theory of relativity; the design and construction of large accelerators would be impossible without this theory.

Today, particle accelerators are as important to nuclear research as microscopes are to bacteriology.

take various forms: a gas cylinder, a metallic filament that emits electrons, or a crucible in which an element that is solid at ordinary temperatures is boiled. In this example, the reservoir is a gas cylinder.

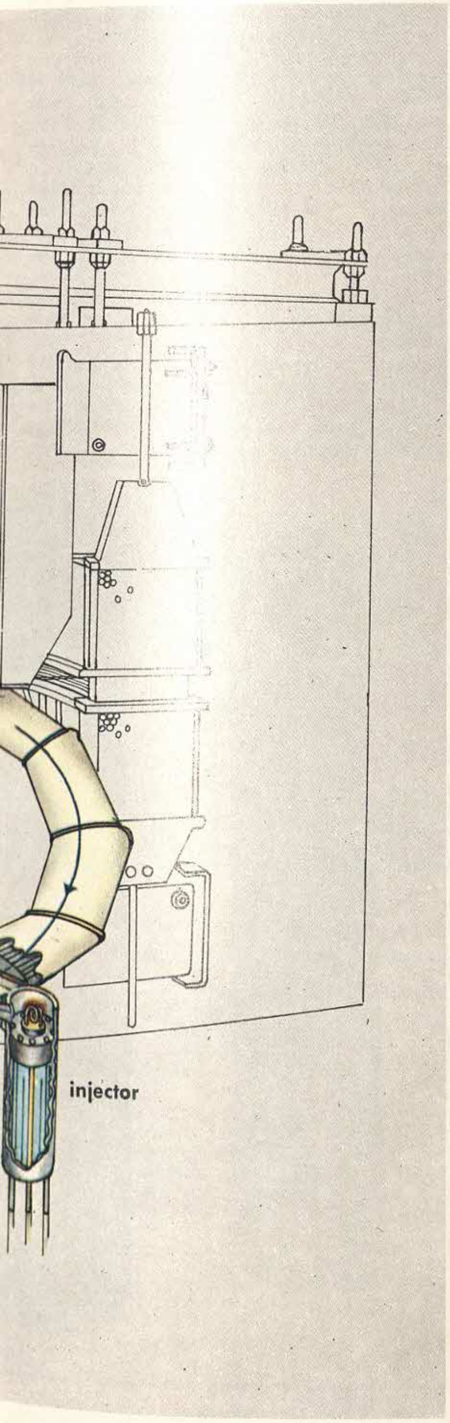
The injector introduces the beam of particles into the accelerator. Atoms in the gaseous state cannot be accelerated because they are neutral and, therefore, are unaffected by electric fields. They must first be ionized by bombardment with a beam of electrons. On the other hand, this preliminary operation is not necessary when electrons are to be accelerated, because these particles are already charged. Finally, the ions or the simple particles must be introduced into the actual acceleration area where they are brought up to the desired speed. Preacceleration is needed to ensure that the particles arrive at exactly the point where they have to begin their acceleration trajectory.

The accelerator is the device that creates the accelerating electric field and determines the name of the entire system. The term *electrostatic accelerator* denotes a device in which the particles are accelerated by means of a constant electric field created between the plates of a condenser. The particles are generally forced to pass several times through the same electric field; their speed is increased at each passage until they reach the desired value.

Auxiliary services are then brought into play. The space in which the acceleration occurs must be completely evacuated. The particles must follow a trajectory that is determined by the characteristics of the electric and magnetic fields that have been purposely created inside the accelerator. If any gas is left in this space, the particles may collide with its atoms and be diverted from their paths, thereby reducing the efficiency of the accelerator. Consequently, the residual gas pressure inside the accelerator must be kept below 10^{-5} mm of mercury. At this pressure the density of the air is about 100 million times less than it is at sea level, and the average distance between the molecules is on the order of several yards. Under these conditions, the probability of collisions between ions and molecules of air is quite small.

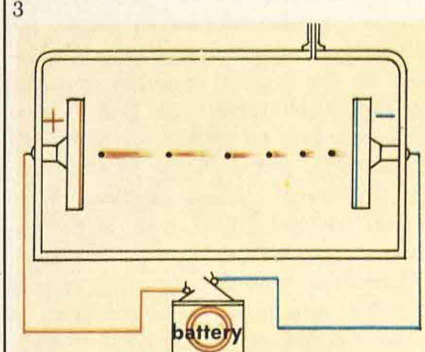
After the particles have been accelerated, they must be extracted from the system and conducted to the place where they will be made to collide with other particles to produce the required impact phenomena. They must be shunted out of the trajectory they are following in the accelerator. In many cases this trajectory takes the form of a circular orbit (hence, the name *cyclotron*). The deviation in the desired direction can be obtained by means of a magnetic field that is positioned along the trajectory and brought into operation at the moment when the particles have attained the desired speed or energy to set up the accelerating fields.

The particles then pass through the extraction window—the "exit door" for the particles. It consists of a thin metallic lamina that the particles must cross to enter the area in which the experiments are performed. These experiments, of course, are widely varied.



THE ELECTRON VOLT—The unit used in nuclear physics for the measurement of energies is the electron volt, usually abbreviated eV. One eV is the energy that an electron acquires when it is accelerated by a potential difference of one volt.

3



NUCLEAR TECHNOLOGY—II | advanced nuclear reactors

Atomic reactor technology is still in a dynamic state of development, hence the varied ideas of reactor designers as to the best method of designing a reactor for a given purpose. Much of the development is aimed at putting atomic power on a competitive footing with conventional power—not an easy task because of the cost of nuclear research. However, thermonuclear power is becoming increasingly more feasible. This article examines the worldwide development of reactors for science and industry, including breeder reactors, and then discusses various moderating (neutron control) and cooling systems now in use.

The use of natural uranium—uranium-235 and uranium-238—in graphite-moderated, gas-cooled reactors requires large plants. This usually constitutes an economic disadvantage and governments usually subsidize these plants. Smaller plants can be built, using enriched uranium; however this fuel is more costly than uranium-235 and uranium-238. Generally speaking, the kilowatt cost is lower for power reactors operating on natural uranium.

BREEDER REACTORS

Uranium-238 is called a fertile material because it can be converted into a fissionable substance. Plutonium production in reactors is formed by neutron irradiation of uranium-238. The capacity of advanced reactors to produce plutonium is accelerated in fast reactors—whose objective is to produce fissile plutonium faster than their fissile materials are consumed—hence the name "breeder." These production reactors supply plutonium for other power plants, and are called fast-breeding superconverters. Thus, apart from generating electrical energy, they produce secondary fuel to satisfy the needs of an expanding electronuclear system. In this respect, emphasis is placed on the doubling time—the time needed to accumulate an amount of plutonium equal to the initial charge of fissile material. Currently, a doubling time of between ten and fifteen years is the objective.

Operation with fast neutrons is only possible with highly enriched cores of at least 20 percent uranium-235 or fissile plutonium. Great amounts of energy are released in the restricted volumes within the core, so that heat has to be efficiently removed. In fast reactors, molten sodium is generally used for this purpose. It leaves the core at a temperature of at least 500° C (932° F) and—because of its high boiling point of 880° C (1,616° F)—the cooling system does not need to be pressurized, thus avoiding many construction problems. A 75,000-kilowatt sodium-graphite power facility is located in Hallam, Nebraska.

Unlike the core of a thermal reactor, the core of a fast reactor is not surrounded by a reflector, because it would thermalize the neutrons. In its place is a blanket of uranium impoverished in uranium-235 (and therefore enriched in uranium-238) that is gradually converted into plutonium.

Currently, thermal reactors extract a low percentage of uranium's energy potential. This percentage can be doubled with advanced converters, although even this value is still low. Superconverters should eventually be able to make thermal reactors profitable. It is commonly agreed that superconverter fast reactors will reach full industrial maturity around 1980. Hopefully, this objective will achieve satisfactory exploitation of nuclear fuel.

COOLING SYSTEMS

Natural uranium is particularly convenient for use in reactors moderated with heavy water. Various cooling systems are possible, but the heavy-water type is the only one widely used on an industrial scale. Two examples of the Canadian type (CANDU—Canadian Deuterium Uranium) are described in Illustrations 1 and 2.

THE PRESSURIZED-TUBE SYSTEM

The pressurized-tube reactor system is common in the United States and is becoming common elsewhere. The system of pressurized tubes adopted for CANDU is not the only one possible for heavy-water systems. Reactors of this type can also be built with a pressure vessel system for the ordinary pressurized water reactor (PWR). This is the design that has been followed for the 50-megawatt

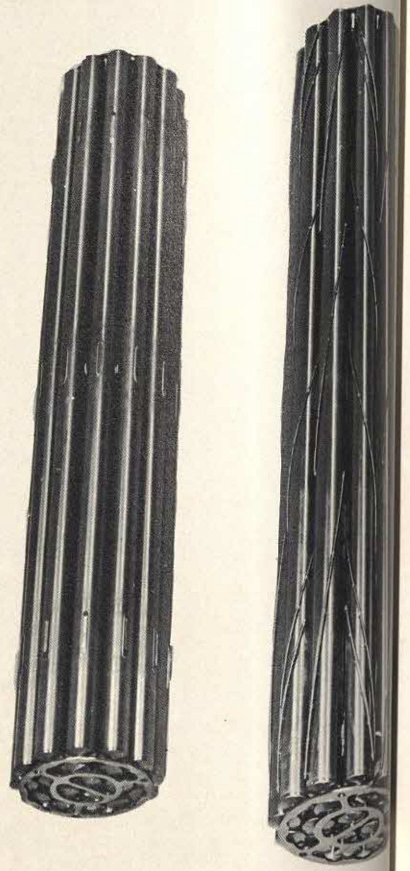
1

CORE OF A CANDU REACTOR—In this reactor the heavy water, which acts as a moderator, is kept separate from the water used as a coolant. The coolant passes through the reactor's core in a series of pressurized pipes arranged in parallel (Illustration 1a). Each pipe is surrounded by a concentric insulation pipe so that the heavy water can be kept at a much lower temperature than the coolant water. The coolant leaves the core at 290° C (554° F); the heavy-water temperature is less than 70° C (158° F), thus avoiding high pressures in the heavy-water vessel. Illustration 1a shows how the pipes are arranged horizontally.

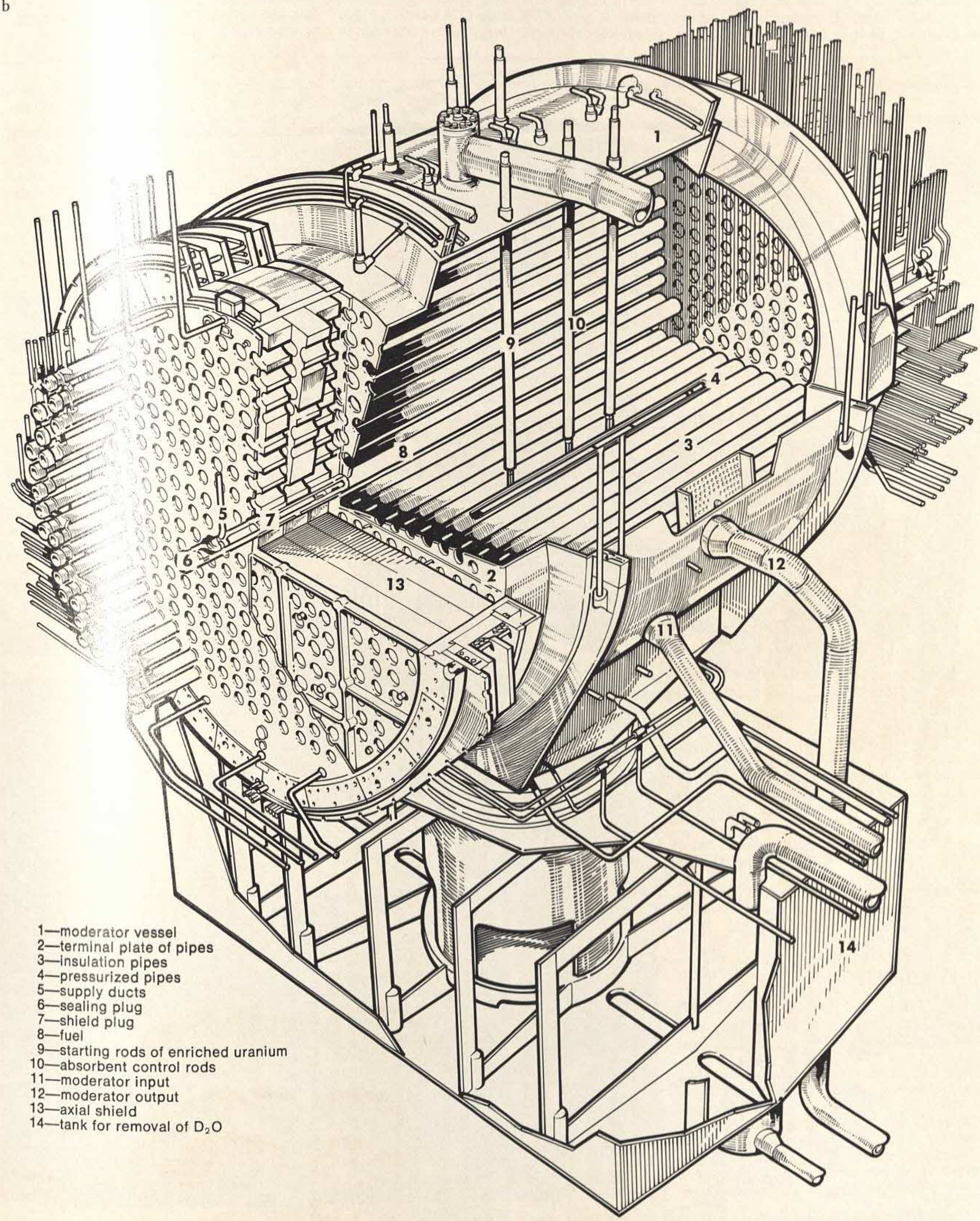
The fuel elements are placed in pressure pipes one after the other. They consist of bundles of 19 rods of natural uranium oxide clad with a zirconium alloy (Illustration 1b). Each bundle is 50 cm (about 20 in.) long. Fuel exchange is continuous while the reactor is in operation. While the use of zirconium alloys (with their low neutron absorption rate) is optional for enriched uranium reactors, it is essential for reactors operating on natural uranium as fuel.

Control of the reactor's power is achieved by means of absorbent rods arranged vertically, and by variations in the moderator level. The reactor is stopped by discharging the moderator into a container located beneath it. It is started by increasing reactivity with the insertion of vertical rods of uranium greatly enriched with uranium-235.

a



b

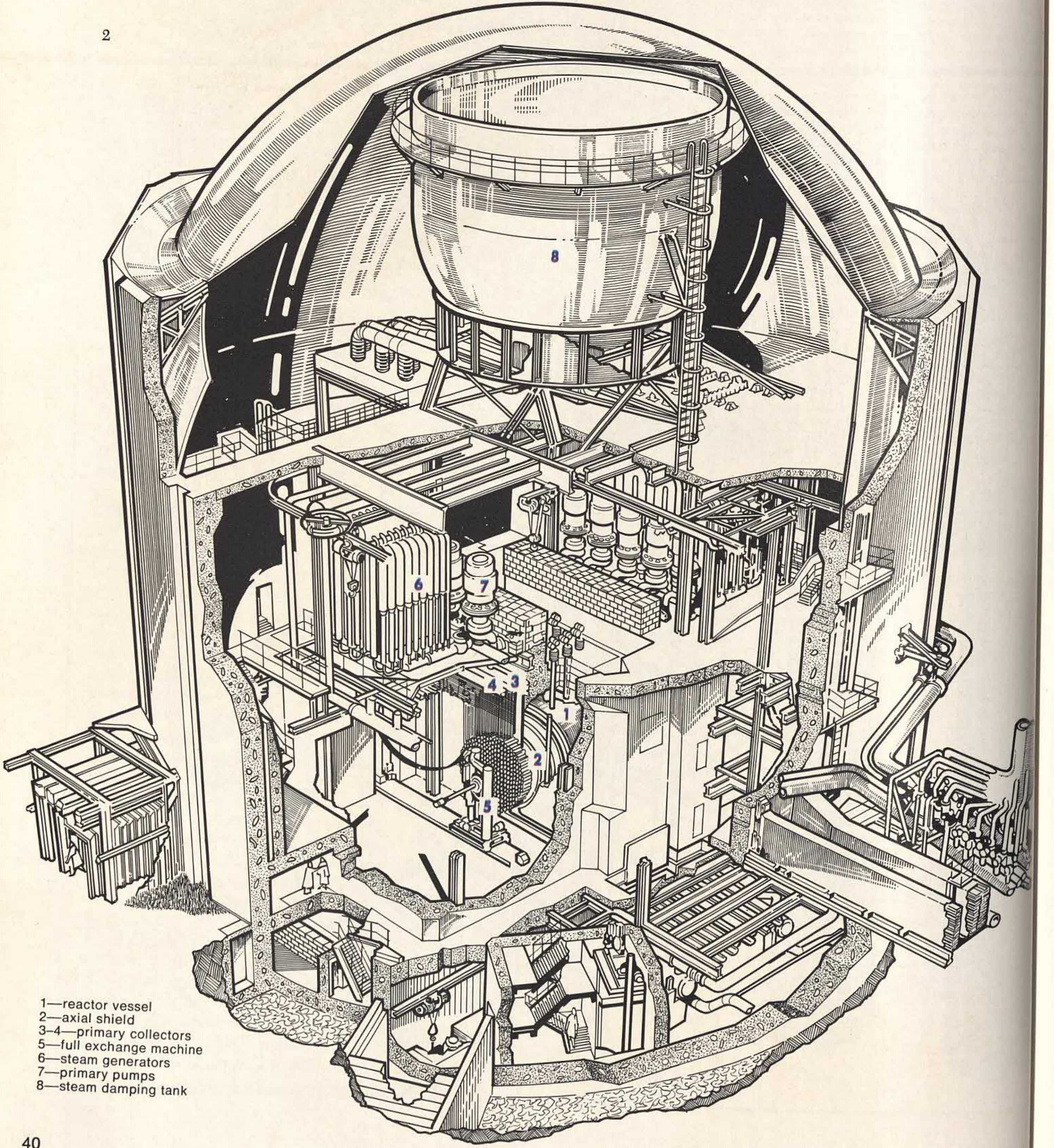


1—moderator vessel
2—terminal plate of pipes
3—insulation pipes
4—pressurized pipes
5—supply ducts
6—sealing plug
7—shield plug
8—fuel
9—starting rods of enriched uranium
10—absorbent control rods
11—moderator input
12—moderator output
13—axial shield
14—tank for removal of D_2O

CROSS SECTION OF POWER PLANT WITH CANDU REACTOR—The reactor is located in the lower part of the building, which is sealed for safety reasons. The large tank above contains enough water to damp down steam that would escape if an accident occurred. As in all pressurized water systems, an indirect cycle is adopted. Therefore, heat exchangers are needed to generate steam for supplying the turbo-generator. The efficiency of these plants is around 29 percent. In Canada, a 200 electrical megawatt prototype is located at Douglas Point on Lake Hudson. Four 500-megawatt units are at Pickering, Ontario.

Even if the reactor moderated with heavy water and cooled with pressurized heavy water has still not reached full industrial maturity, it is approaching it very quickly.

2

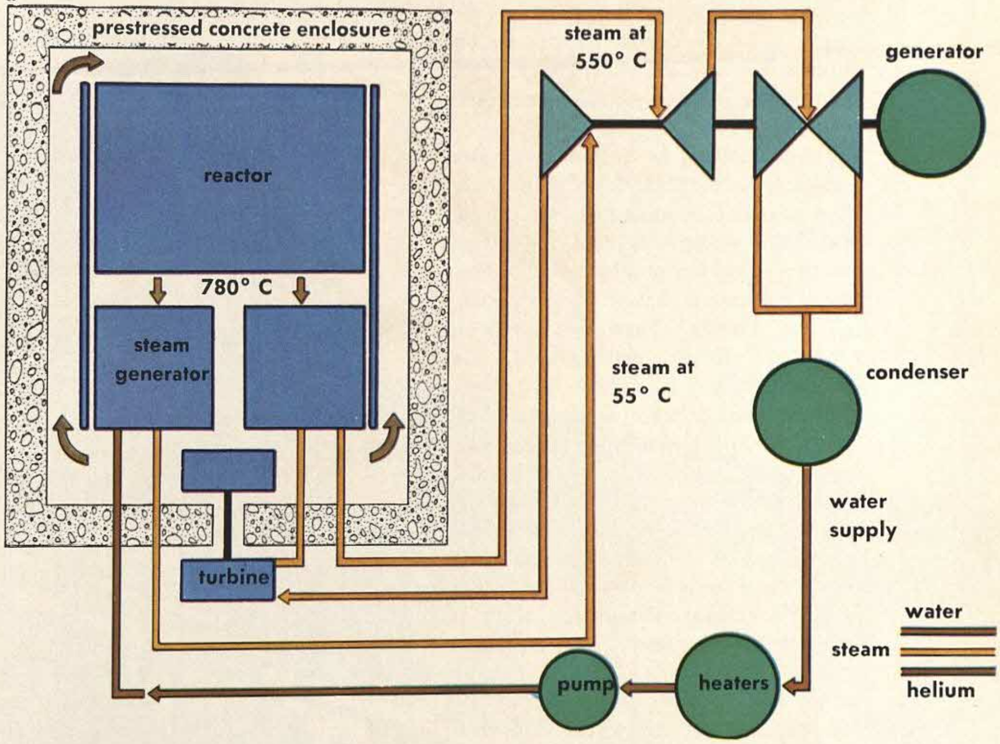


1—reactor vessel
2—axial shield
3–4—primary collectors
5—full exchange machine
6—steam generators
7—primary pumps
8—steam damping tank

3

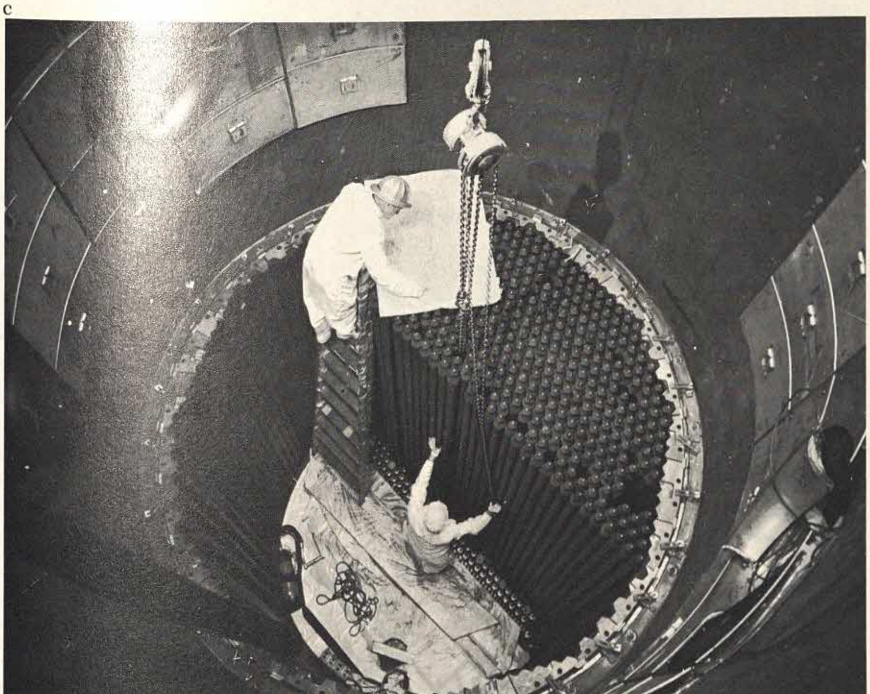
A PLANT WITH AN HTGR REACTOR—This graphite-moderated reactor is cooled with high-temperature gas. To extract the heat from the core, helium is used rather than the carbon dioxide of ordinary gas reactors. The chemical inertness of helium allows work at extremely high temperatures without causing any deterioration in the materials that it contacts. In the coolant leaving the core, temperatures can thus go as high as 800° C (1,472° F).

a



b

c



In an indirect cycle with a secondary steam phase, efficiency can reach 40 percent. Illustration 3a shows a plant layout with an HTGR reactor. If the temperature is raised still further to 1,000° C (1,832° F), it is possible to adopt a direct-cycle layout with a gas turbine, thus achieving an efficiency of around 50 percent. For such high efficiencies, core materials must have high resistance to chemical attack and thermomechanical stress. The chemical attacks come from impurities in the helium. Since metals cannot be used for shielding the fuel elements, a special type of graphite has been adopted. It is given a special treatment to make it impermeable to the fission products. Moreover, graphite is fairly transparent to neutrons. Fuel elements clad in graphite of various forms contain enriched uranium oxide, sometimes to the extent of 90 percent in uranium-235.

In Great Britain, the Dragon, a low-power experimental reactor of 20 thermal megawatts, has been operating for some years. In the United States, an HTGR prototype supplying a 40-megawatt power station has been operating for several years at Peach Bottom, Pennsylvania. Illustration 3c shows the inside of this reactor while its more than 800 fuel elements are being loaded. The assembly of one of these elements is shown in Illustration 3b. Fuel rings of sintered oxide are being threaded onto a graphite rod inside a shield that is also made of graphite.

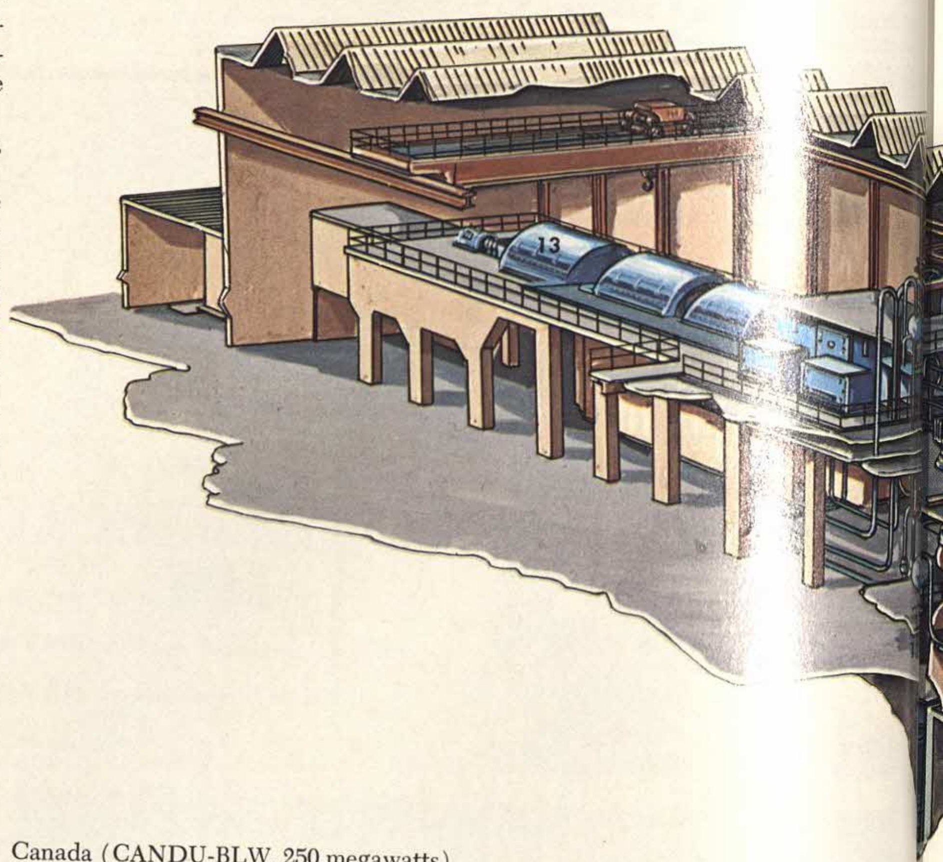
MZFR prototype at Karlsruhe, Germany.

To achieve a high degree of fuel exploitation, the use of natural uranium means a continuous exchange of fuel elements. In PWRs, the exchange is discontinuous. The mechanical systems of exchange through the pressure vessel are naturally rather complex.

Besides cooling by pressurized water, organic fluids (diphenyls and terphenyls) have also been considered. An organic cooled and moderated plant (11,000 kilowatts) is located at Piqua, Ohio. Four power prototypes based on gas cooling (carbon dioxide) have been built in France, Germany, Switzerland, and Czechoslovakia. This type of reactor tends to be uneconomical because of the difficulties involved in making fuel cladding that is simultaneously resistant to high temperatures and fairly transparent to neutrons.

Cooling with boiling water led to the construction of some pilot plants. One of these—the power station at Marviken, Sweden (200 megawatts)—has a pressure vessel system. In this system, the boiling heavy water acts as both moderator and coolant. Called a heavy-water BWR (boiling water reactor), it has a direct cycle so that the alternator turbine is driven by steam produced from heavy water. Of course, every part of the plant—including the conventional plant—must be completely gas-tight to avoid losing such an expensive substance as heavy water. Several boiling-water reactors are located in the United States. The Pathfinder Atomic Power Plant in Sioux Falls, South Dakota, is a boiling-water type with integral nuclear superheating.

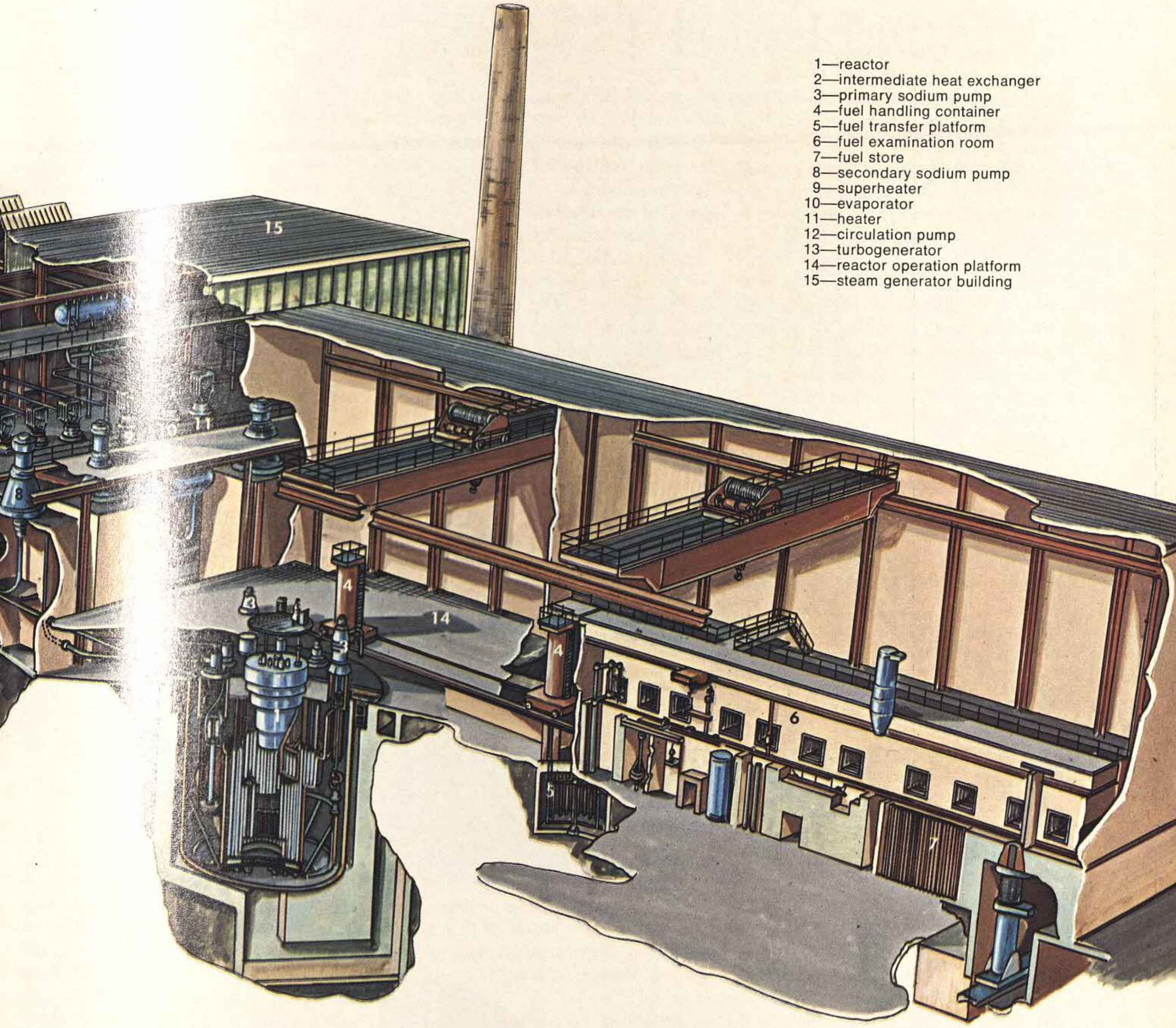
The other version of direct-cycle cooling with boiling water is based on the pressure pipe system. Ordinary water boils inside the pipes containing the fuel. The heavy-water moderator surrounds the pressure tubes and is isolated from them thermally by external concentric pipes, as in the CANDU. In this way, the heavy water is at a low temperature, while the high-pressure cooling circuit does not raise special sealing problems because it contains ordinary water. This type of reactor is being developed in Canada (CANDU-BLW, 250 megawatts) and Italy (CIRENE-BLW). In Great Britain, a similar reactor (100 megawatts) has been built, although it operates on an enriched uranium core. A reactor of this type, with a capacity of 17,000 kilowatts, is located in Parr, South Carolina.



One of the chief advantages of heavy-water reactors is that they ensure a high utilization of uranium resources. When one asks how much natural uranium is needed in order to obtain slightly enriched fuel for PWR or BWR reactors by means of isotopic separation processes, it is found that heavy-water reactors produce more electrical energy per unit of weight of natural uranium. This also depends on the fact that heavy-water reactors convert a particularly large amount of uranium-238 into plutonium. This element is largely made up of fissile isotopes and may, therefore, be regarded as a secondary fuel.

The residual plutonium is extracted from the fuel elements at the end of each utilization cycle and used again in the same or a different reactor. For this reason, the most efficient heavy-water reactors are usually classified as advanced converters.

This same category includes another type of reactor of completely different concept: the HTGR (High-Temperature Gas-cooled Reactor). It also converts large amounts of uranium into plutonium, guaranteeing an effective exploitation of primary fuel resources. The HTGR is moderated with graphite and cooled with high temperature gas.



1—reactor
2—intermediate heat exchanger
3—primary sodium pump
4—fuel handling container
5—fuel transfer platform
6—fuel examination room
7—fuel store
8—secondary sodium pump
9—superheater
10—evaporator
11—heater
12—circulation pump
13—turbogenerator
14—reactor operation platform
15—steam generator building

FAST REACTOR—These reactors are being developed in various countries, especially the United States, Great Britain, the Soviet Union, and the Euratom countries. The first fast reactors built were used by the United States to power nuclear submarines. Brilliant solutions were found for exceptional operating conditions, but the solutions were uneconomic.

Important plants are located in Great Britain and the Soviet Union. The British PFR plant (Prototype Fast Reactor) shown in cross section here is at Dounreay, Scotland. This plant generates a power of 750 electrical megawatts. The Soviet prototype, which is called BN-350 (from the Russian, *Bystrye Neitrony*—"fast neturons"), is located in Shevchenko.

The reactor shown is capable of supplying a 150-megawatt power station and also, by providing energy to a saltwater distillation plant, can produce 120,000 m^3 (about 32,000,000 gal) of drinking water per day.

PETROLEUM—I | from microorganisms to petrochemistry

Petroleum, one of the most useful raw materials known to man, is a mixture of gaseous, liquid, and solid hydrocarbons found in varying amounts in sedimentary rock deposits throughout the world. In its crude state, petroleum is practically useless. When refined, however, it supplies fuels, lubricants, illuminants, and surfacing materials. Important products derived from petroleum include antiseptics, drugs, cosmetics, solvents, paints, detergents, plastics, synthetic rubber, and fibers. Today, petroleum is generally called oil. The term crude oil refers to petroleum as it comes from the Earth.

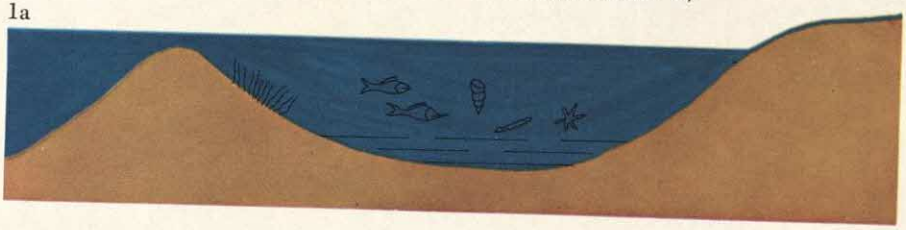
Petroleum has been exploited in various forms since remote times. The ancient Egyptians coated mummies with asphalt, which is a semisolid form of petroleum. The Chinese used natural gas for fuel as long ago as 1,000 B.C. About 600 B.C., in Babylonia, asphalt was used as a bonding and waterproofing material for walls and as a paving for streets. American Indians used crude oil as a medicine hundreds of years before the arrival of the white man.

Two important factors, both connected with the Industrial Revolution, led to the birth and rapid growth of the petroleum industry. The first was a need for a better and cheaper source of lighting for homes and factories; the second was a need for lubricants for the new machines.

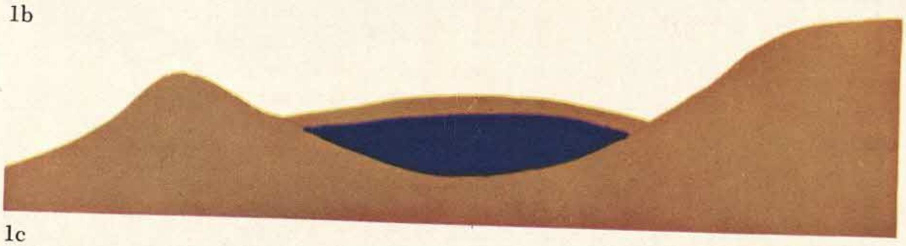
In 1859, on land near Titusville, Pennsylvania, Edwin L. Drake drilled the first oil well. He had devised methods and machinery to do the work rapidly, and introduced the use of iron pipe to prevent cave-ins. His success marked the beginning of the modern petroleum industry.

THE ORIGINS OF PETROLEUM—Millions of organisms live in the warm surface waters of shallow seas (Illustration 1a). When they die, they fall to the sea floor and mix with the mud and sand deposited there. Later, the sea recedes and new rock formations bury the old sea floor (Illustration 1b). As time passes, processes that are still not clearly understood change the buried organic matter into hydrocarbons (Illustration 1c).

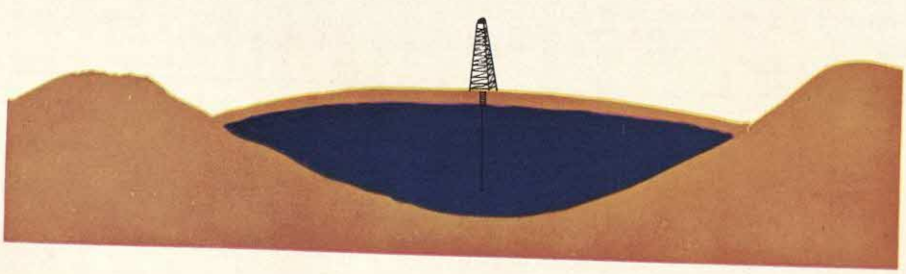
1a



1b



1c



HOW PETROLEUM WAS FORMED

No one knows for sure how petroleum is formed, but most scientists accept the organic theory, which concludes that it was formed hundreds of millions of years ago from the remains of small aquatic plants and animals when the sea covered large areas of what are now landmasses. As the animals and plants died, they sank to the seabed where they mixed with mud and sand in layers called marine sediments. Later, these sedimentary layers were covered by more mud and sand, which finally turned into rock. As centuries passed the sea withdrew. The crust of the Earth heaved and buckled, burying the sedimentary deposits; the heat and pressure caused by the overlying rock, together with decomposition of the organic life, formed oil from the animals and plants in the deep-buried layers. Petroleum is thus called a fossil fuel because, like natural gas and coal, it is derived from organic material. It is also referred to as a mineral oil since it lies in the pores of underground sedimentary rocks or sand and is linked with the minerals.

Chemically, the major components of crude petroleum are hydrocarbons, compounds of carbon and hydrogen, combined in molecules of different sizes and arrangements. Small molecules (one to four carbon atoms) constitute the gases; larger molecules (from four to about ten carbon atoms) constitute gasoline; still larger molecules (up to 50 atoms) make up the light fuels and lubricating oils; and gigantic molecules (up to several hundred carbon atoms) constitute the heavy fuels, waxes, and asphalts. Depending on origin, some crude oils are rich in gasoline components, some in kerosene, and others in oil and waxes.

WHERE PETROLEUM IS FOUND

At the beginning of the twentieth century, geologists began to map land features as indicators of favorable places to drill for oil. Particularly promising were outcroppings that gave evidence

2

WORLD DISTRIBUTION OF PETROLEUM— The areas in blue are zones rich in petroleum. Once these regions were completely covered by shallow seas filled with plants and organisms that, over the centuries, formed layers of material that evolved into hydrocarbons.

of alternating layers of porous and impermeable sedimentary rock. Porous rock, typically sandstone, limestone, or dolomite, acts as a reservoir for any petroleum present. In porous rock, petroleum can move, under a pressure differential, through the interstices and cracks to a point of withdrawal, that is, to the bottom of a well. Impermeable rock, typically clay or shale, acts as a trap, preventing the migration of petroleum from the reservoir. The probability of discovering exploitable amounts of oil is highest when the alternating strata are uplifted into domes or anticlines, for under the domes the impermeable or "cap" rock traps all the petroleum that migrates from more distant reservoirs. The thicker the reservoir—the range is from a few feet to several hundred—the greater the potential yield of oil; the more permeable the rock, the greater the fraction of the oil trapped in the reservoir that can be recovered.

Petroleum is found in two types of oil fields: primary and secondary. Primary fields are formed of mother rocks consisting of layers of marine microorganisms and detritus deposited by the sea. Generally these rocks are claylike silicas and limestone that do not contain the usual calcareous skeletons, despite the fact that they were dissolved by salt water from the sea bottom. In these fields, petroleum is found where it originated, in layers that once formed the bottom of the sea.

Secondary oil fields are made up of younger rocks to which petroleum migrated. Rocks in these fields are rather porous and permeable. They have trapped and accumulated hydrocarbons enclosed by impermeable rocks. The hydrocarbons left the mother rocks and traveled through porous and permeable layers until they were blocked by an impermeable layer. The petroleum then accumulated in front of the obstacle, forming a secondary oil field.

TYPES OF OIL FIELDS

Petroleum deposits can also be classified according to the type of fluids they contain. The three main types are oil, dry gas, and condensed gas. Three conditions determine the state of the hydrocarbons

in a petroleum deposit: (1) the chemical composition of the mixture; (2) the pressure condition of the oil field; and (3) the temperature of the oil field. An oil field contains a large quantity of hydrocarbons. Generally such a field also contains gas and salt water.

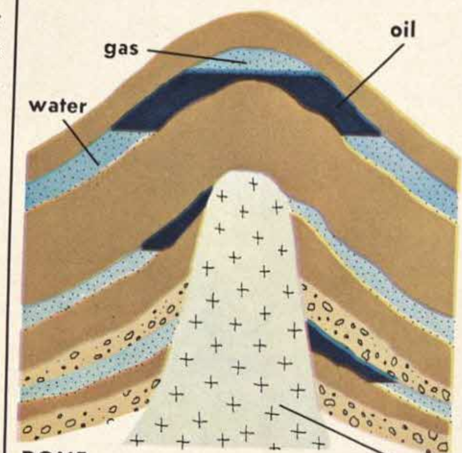
Because of differences in specific gravity, the fluids are distributed so that gas is on top, oil in the middle, and water underneath. This kind of field can be without free gas, so called to distinguish it from the gas that is always found dissolved in petroleum.

Gas fields may be of dry gas or condensed gas; such a distinction refers chiefly to the composition of the hydrocarbon mixture. In the first case the mixture is composed almost exclusively of methane, and it is called dry gas because it is difficult to condense and liquefy. Certain pressure conditions can liberate the gas from the oil, thereby producing a change from liquid to gas. The reverse situation can also take place, with gas changing to liquid, but this occurs only under certain temperature and pressure conditions. With methane these conditions are never encountered during the passage from the bottom of the well to its ultimate destination; as a result, methane always remains in the form of gas. The methane oil fields are dry gas hydrocarbons.

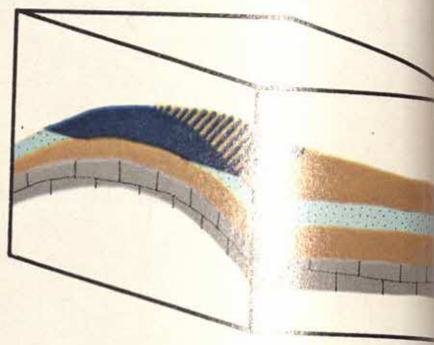
In condensed gas fields the hydrocarbon mixture is formed of methane and higher hydrocarbons such as ethane, propane, and butane. These mixtures, different from the preceding ones, can meet with pressure and temperature conditions that liquefy some of their components and form gasoline during passage from the well to the surface. The presence of hydrocarbons in the subsoil can sometimes be detected by indications in the covering rock, such as traces of oil with iridescent reflections on moving water, sulfuric salt water, traces of inflammable gas, and small volcanoes of mud with escaping gas.

3

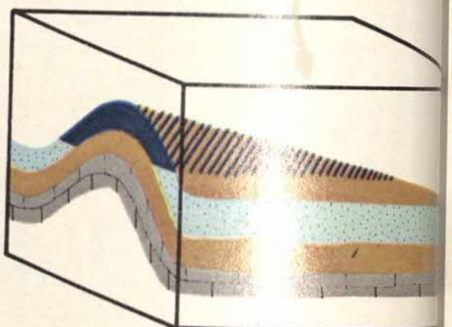
DOMES AND ANTICLINES—Oil fields are often found over salt domes, where rock layers, folded in the shape of a helmet, have been deformed by the formation of salt that pushed them upward. Oil fields also accumulate in anticlines. The hydrocarbons move upward through porous layers and are trapped by an impermeable layer covering the oil field itself. Within the porous layers water is found on the bottom, oil in the middle, and gas on top.



DOME



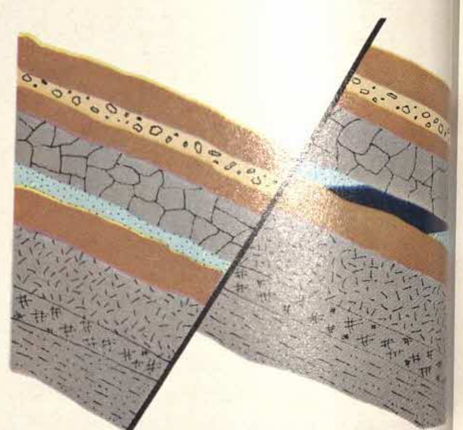
DOME



ANTICLINE

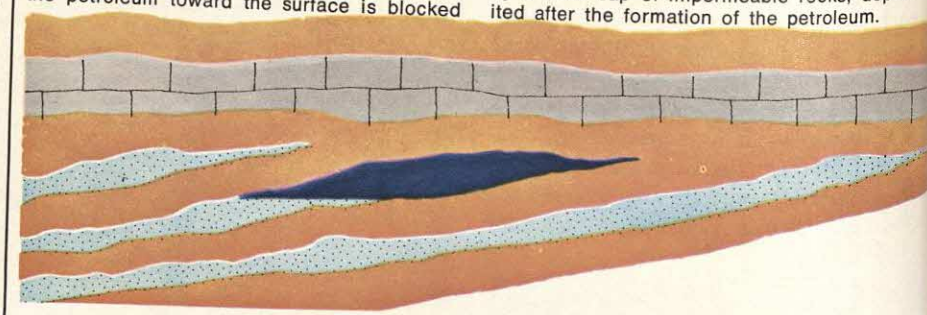
4

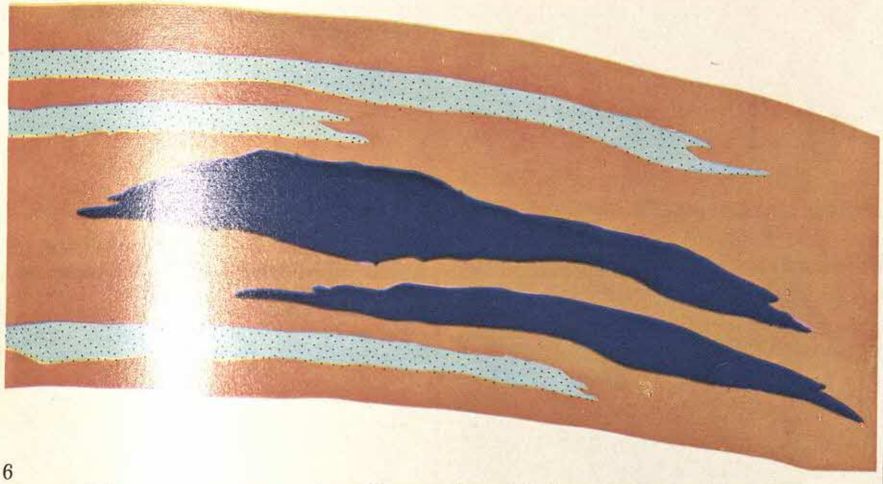
FAULTS—When a complex of rock layers breaks in two due to internal forces, and the parts shift vertically with respect to one another, they form a geological structure called a fault. To picture this phenomenon, imagine three or four sheets of steel placed in a shearing machine. As the sheets are cut, the part resting on the base of the machine remains there, whereas the part under the blade is displaced downward. If the machine is stopped before it has finished cutting, one part of the steel will be lower than the other. When rocks shift in this way, oil may be trapped in porous layers moved against impermeable layers.



5

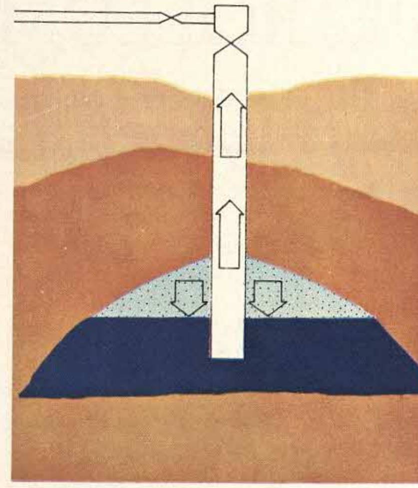
IMPERMEABLE ROCKS—The movement of the petroleum toward the surface is blocked by a thick cap of impermeable rocks, deposited after the formation of the petroleum.





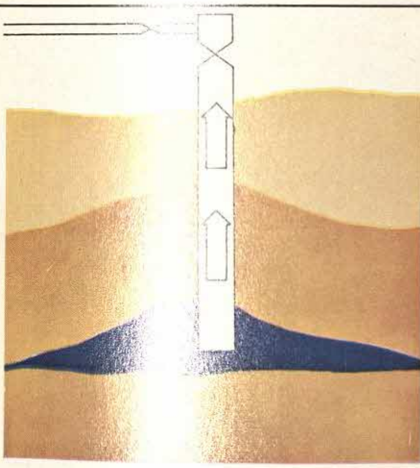
6

LENSES—When certain areas of sand are completely enclosed by a thick, impermeable layer, oil fields are formed that, because of their characteristic shapes, are called lenses.



7

8

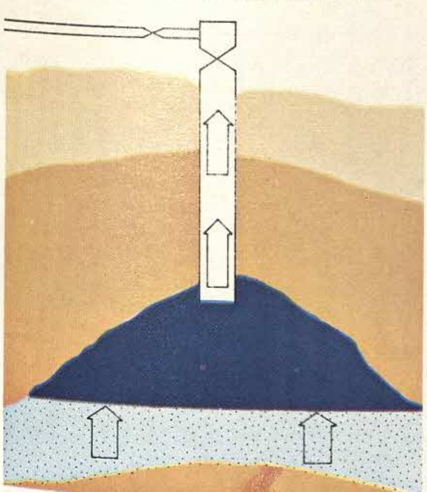


AN OIL FIELD DELIVERED BY DISSOLVED GAS—This type of field contains no water underneath nor gas above the oil, but there is gas dissolved in the petroleum itself that triggers the delivery of the oil. Such a field functions like a bottle of champagne. When the cork is loosened, gas dissolved in the wine (oil) pops it out. An oil field driven by dissolved gas is the least efficient type because only a small amount of petroleum at the bottom can be recovered. For every 1,000 liters (about 6 barrels) of oil, only 200 to 300 liters (about 1 to 2 barrels) reach the top. The remainder stays on the bottom because the pressure decreases rapidly as the gas escapes. As soon as the pressure lowers a little during delivery the gas is liberated and no more remains at the bottom to push the petroleum to the surface.

AN OIL FIELD DRIVEN BY GAS FROM THE CAP—In this type of oil field, the oil and gas are separated, with the oil at the bottom and the gas on top. The pressure of the mantle or cap of gas pushes the oil to the surface. For every 1,000 liters (about 6 barrels) of oil at the bottom, 300 to 800 liters are recovered, depending on the volume of the cap. The greater its volume, the longer the duration of the drive it supplies and, therefore, the greater the amount of oil driven toward the surface. In such fields the well must draw only from the mineralized layer of oil. Only oil must be extracted in order not to diminish the volume of the gaseous cap. When the layer is exhausted, and no more oil can be extracted, the gas can be exploited.

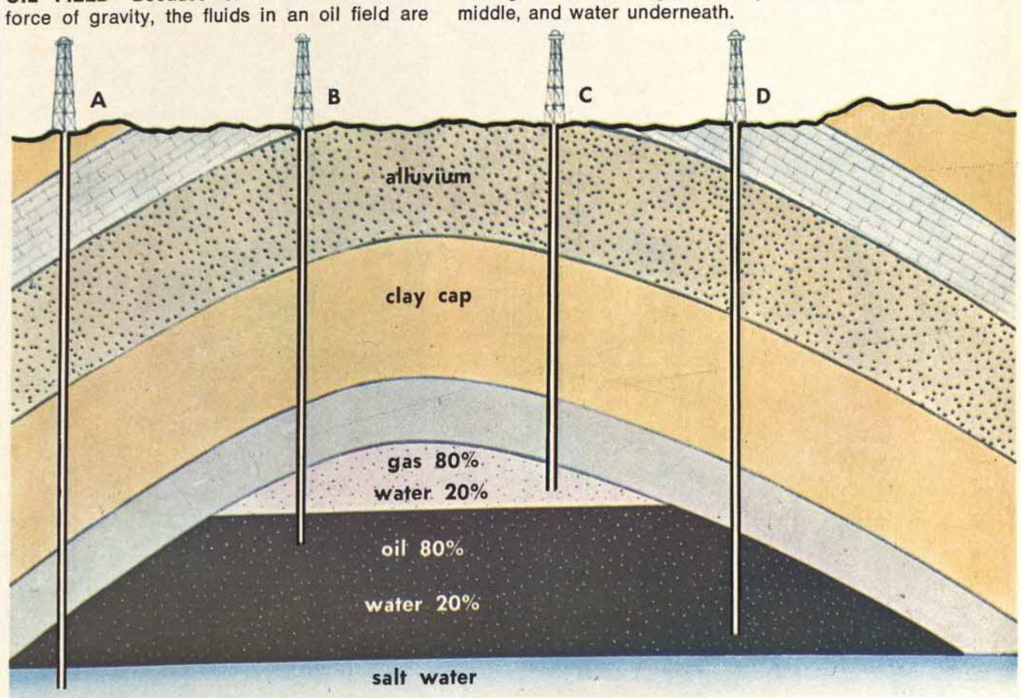
9

AN OIL FIELD DRIVEN BY WATER—When petroleum and water exist under pressure in a layer, the water pushes the petroleum up and out from below. For every 1,000 liters of oil on the bottom, about 850 liters (about 5 barrels) are recovered. During delivery the water rises and fills up the space left free by the petroleum, continuing to supply the necessary drive for delivery. The pressure remains constant if underground passages connect the reservoir with surface waters.



10

THE RELATIVE POSITION OF FLUIDS IN AN OIL FIELD—Because of the effect of the force of gravity, the fluids in an oil field are arranged according to their respective specific gravities, with gas on top, oil in the middle, and water underneath.



PETROLEUM—II | the search for oil

Petroleum, which comes from the Earth as a dark liquid, has been called "black gold" because of its value to mankind. In the early days of the petroleum industry, finding oil was mostly a matter of luck. Early oilmen made and lost fortunes overnight. Even today, locating is a gamble, because no physical or chemical property of petroleum has been found that enables it to be detected from the surface. Thus the exploratory techniques —geological, geophysical, and geochemical—that have been developed do not yield perfectly accurate forecasts of discovery; at best they give the geologist only an indication of how good his chances are. Only by drilling will he learn if oil actually exists in a given location.

Drilling an oil well may cost from $100,000 to $2 million, depending on its depth, location, and other conditions. Thus the preliminary work of geologists and geophysicists, though not always exact, is essential for a successful venture.

While scientists have no way of determining exactly where petroleum has formed underground, they do know the kinds of rock formations that are likely to contain oil.

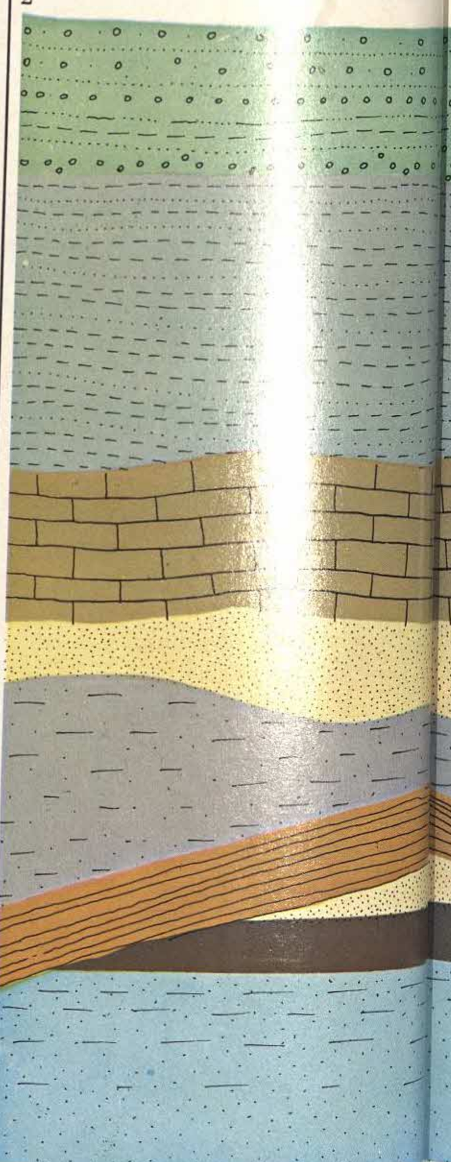
The search for oil begins with the geologist, who has a broad knowledge of the Earth's crust and its history. He focuses his attention on areas that were once the floors of shallow seas where sedimentary rocks are found—sandstone, limestone, and shale—rocks most often associated with petroleum deposits. The paleontologist may be called on to identify fossil flora or fauna found at different levels, as a means of dating the rock strata and determining under what conditions the rock was formed. The geochemist may analyze rock samples, determining chemical and mineral content and electrical conductivity, to see if they are similar to rocks associated with petroleum in other parts of the world.

Information about rock types and structure deep underground is provided by the geophysicist. The tools he uses include the gravimeter, which measures differences in the pull of gravity caused by different densities of underlying rocks; the seismograph, which measures the time needed for shock waves to travel underground; and the magnetometer, which records differences in the Earth's magnetism due to varying rock strata.

Preliminary research along these lines should establish, before drilling is considered, that underlying rock strata were formed under conditions favorable to the formation of oil, and that there are good geological reasons for believing that underground strata contain oil.

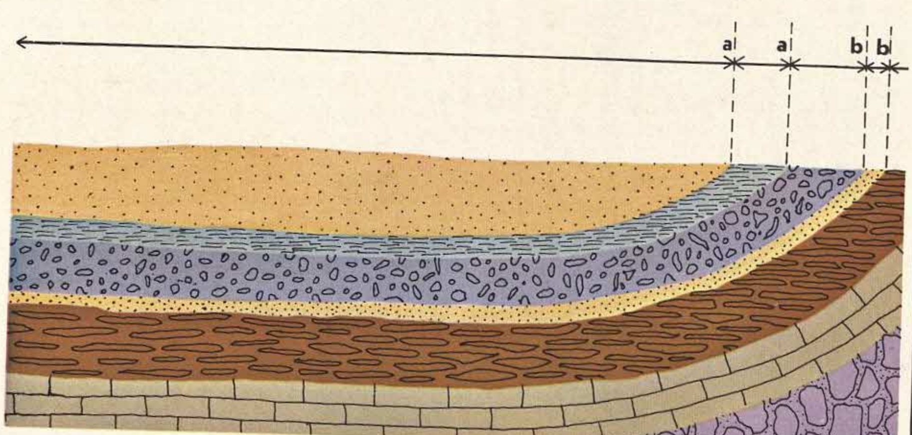
THE STRATIGRAPHIC W... —Geological research provides informat... nly about rock strata exposed or near th... face. Often the places where the strata ar... posed lie a long distance from a suitable ... e for drilling a well. In this case, geolog... search may be

2

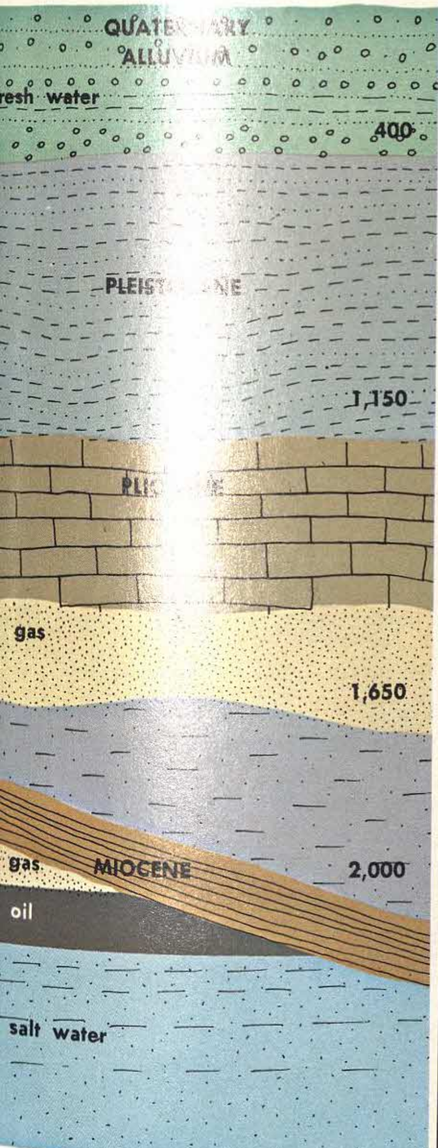


GEOLOGICAL STUDIES OF THE EARTH'S SURFACE—Sedimentary rocks are formed in horizontal layers, one on top of the other, with the oldest at the bottom. A later folding of the Earth's crust, or erosion, may expose a cross section of a layer, as at **aa** in the illustration. Narrow strata may appear in points as at **bb.** The angle at which the strata descend indicates to the geologist where the same strata can be located underground. The succession of different layers is called a stratigraphic sequence.

1



supplemented with stratigraphic wells drilled directly into the rocks suspected of bearing oil. This illustration shows the sequence of layers discovered by a stratigraphic well drilled recently.



3

a

b

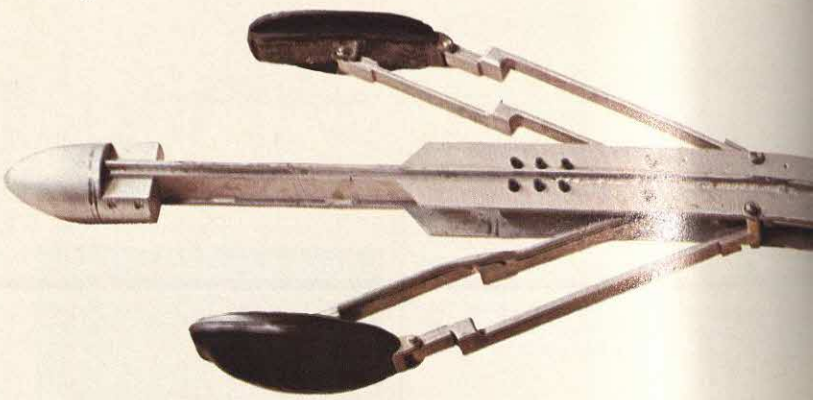
EXPLORING BY MEANS OF THE STRATIGRAPHIC WELL—The well is drilled with a heavy steel bit that bores through solid rock by crushing and grinding it. Liquid mud—called drilling mud—immerses the bit when it is in operation. When this mud is pumped

c

to the surface, it is mixed with bits of crushed rocks. The larger pieces, called cuttings, are examined to identify the strata from which they come. By studying these cuttings, the geologist can establish the depths at which different strata begin and end.

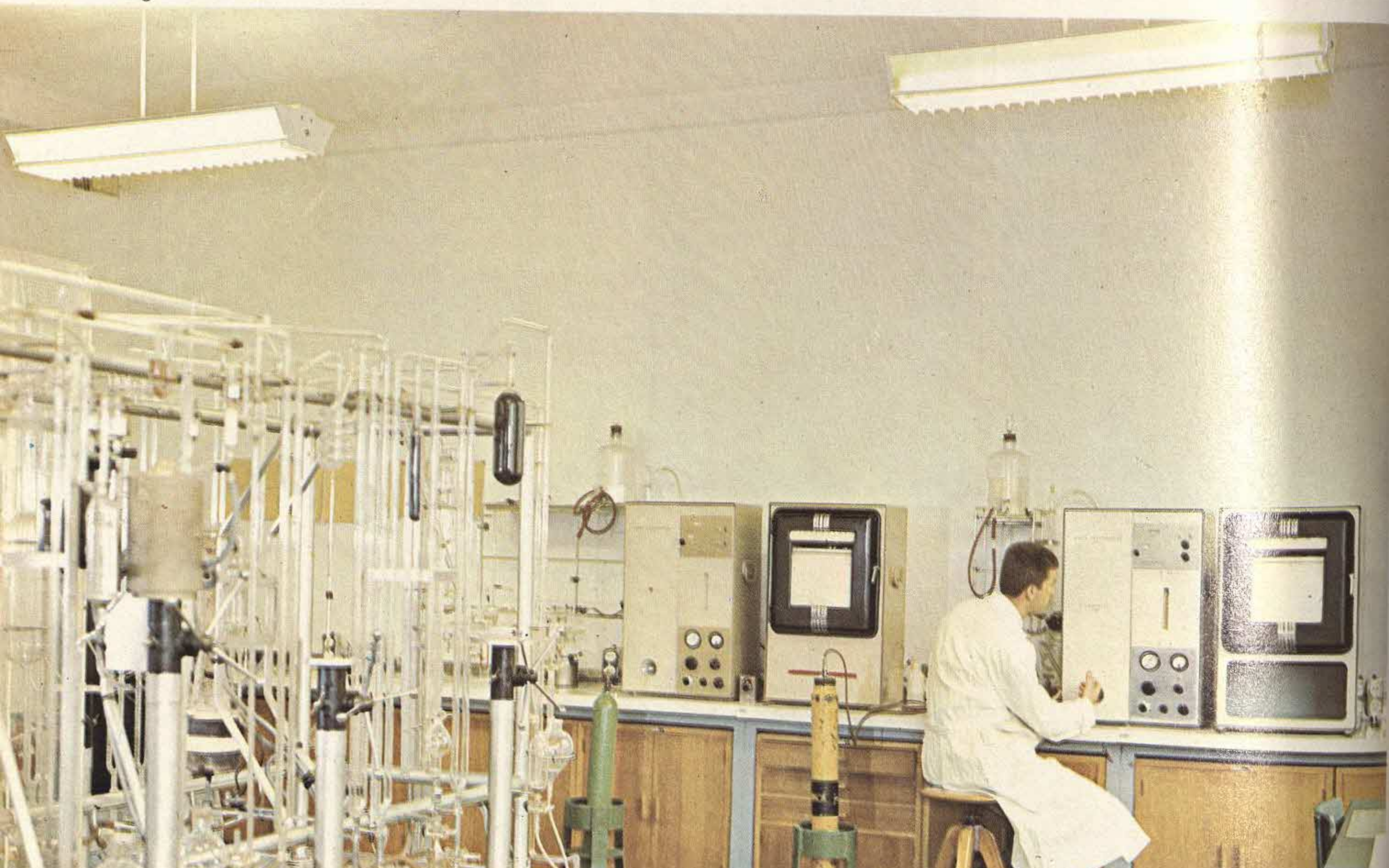
The most useful samples from stratigraphic wells are "cores"—cylinders of rock brought to the surface intact. Illustrations 3a, 3b, and 3c show cores brought up from strata at different depths.

4



CORE BORING—Shown here is the core drill that can be lowered into the stratigraphic well to bring up, intact, a cylindrical sample of the rock or other material from deep underground. This kind of drill provides the samples most useful in determining the physical and chemical properties of rock strata at different levels.

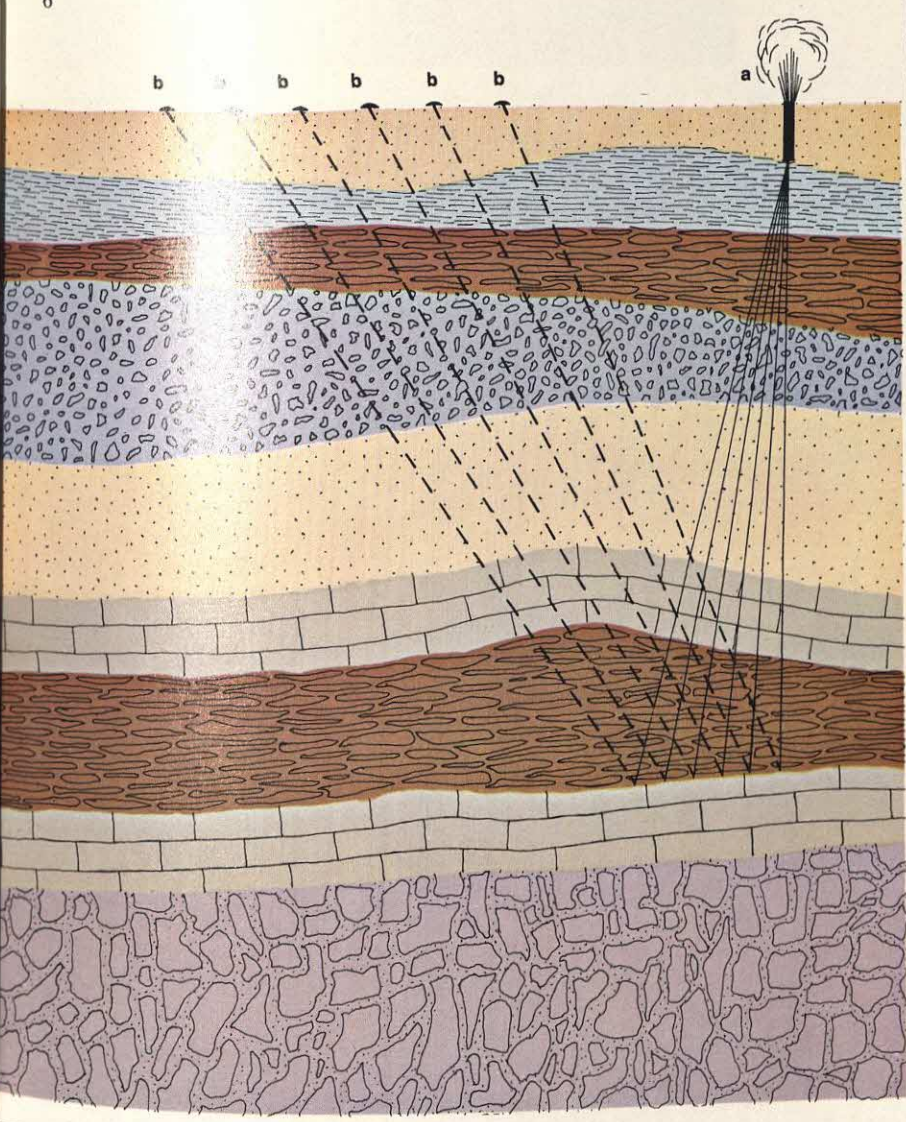
5



GEOPHYSICAL RESEARCH — In this well-equipped laboratory, thorough chemical analyses are made of rock samples from the different levels of stratigraphic wells, to determine if the rocks were formed under conditions favorable to the formation of petroleum.

GEOSEISMIC SURVEY—Once it has been established that a certain stratum is likely to contain oil or natural gas, it is necessary to determine places where there might be an accumulation large enough to justify the high cost of drilling. To map the depth and thickness of the stratum, oilmen bounce seismic waves off the interfaces, which are the boundaries between strata. The seismic wave is produced by an artificial explosion at ground level. In this illustration, an explosive charge detonates at **a** and the waves are picked up at the points **b** by geophones, which are instruments similar to microphones. From the time required for the wave to travel down to the boundary and back to the geophone, the depth can be calculated. This system of exploration makes it possible to map rock formations that lie deep underground or beneath the sea.

6

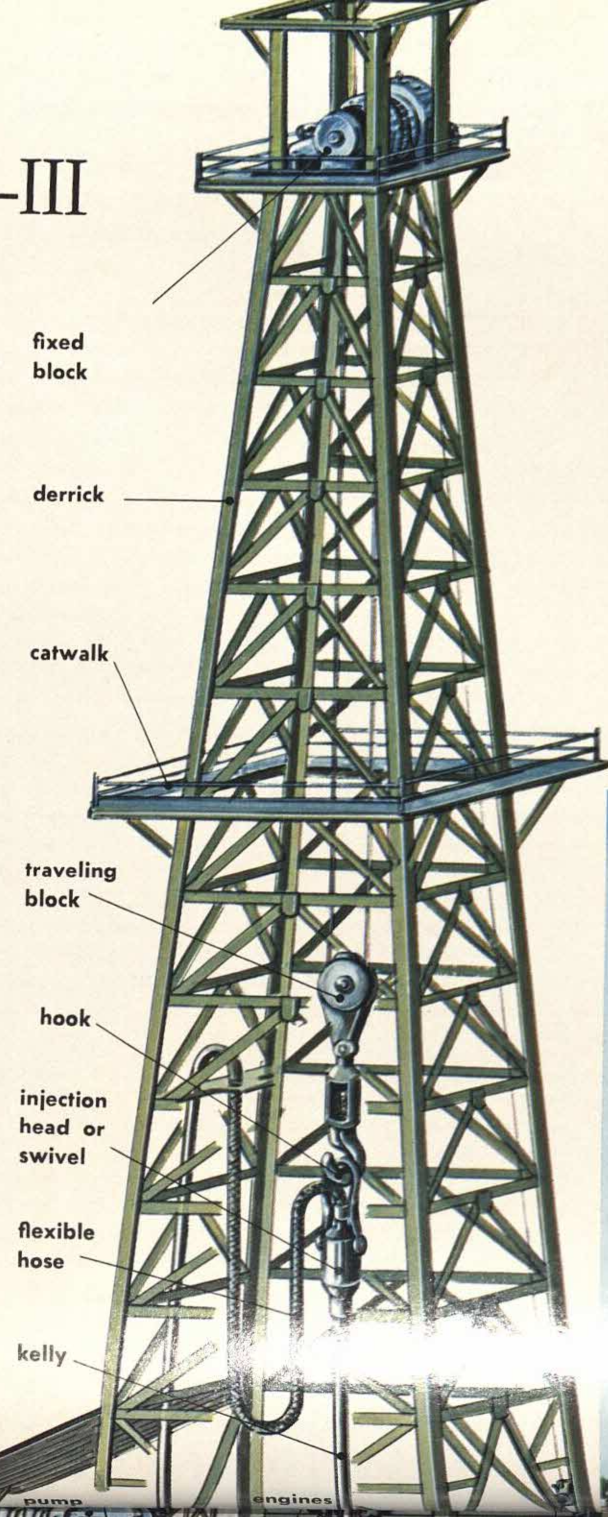


PREPARATION FOR DRILLING—After extensive surveys have located geological formations likely to contain oil, actual drilling begins. It will then be determined if results can match the promises of exacting and costly research.

7

PETROLEUM—III

After geologists and geophysicists have completed their research and concluded that a given site is a probable source of petroleum, the next step is drilling to verify the deposit. The area to be worked is laid out—usually it is a rectangle of 60 to 80 m (about 200 to 260 ft) by



1

THE DRILLING RIG—The drilling rig may be of different sizes and forms, depending on the depth to which it must reach. Some are as tall as 80 m (about 260 ft), with a framework capable of supporting more than 400 tons. A turntable in the center of the derrick floor turns and supports the drill pipes.

At the top of the derrick, the crown, or fixed block, usually of six pulleys, supports a traveling block of five pulleys by a steel cable. The drilling battery of tools and drill pipe is supported and maneuvered by a hook on the traveling block. The battery is composed of the following parts: a swivel or mud injection head; a safety device made up of a series of chokes to contain any gusher that may occur in the shaft; a kelly (a hollow tube with four or six sides), which, fixed in the rotating turntable, drives the entire battery; drill pipes between the kelly and the drill collars, their number depending on the depth of the well (they transmit rotary motion to the bit); drill collars attached to the drill pipes (they concentrate the weight applied to the bit as low as possible in order to avoid any bending of the drill pipes and any consequent deviation in the shaft); and the drilling bit, the shape of which varies according to the hardness of the rock it must penetrate.

No less important are the engine and pump groups. The engines, usually of the internal combustion type, drive the turntable, the winch that raises and lowers the drilling rig, and the pumps. The pumps circulate the drilling mud—a mixture of clay and water with special chemicals added to compensate for varying composition of the water and the formation being drilled—at a high rate of flow and under great pressure. From the feed line, the mud is pumped into the swivel or injection head, down through the stems and out

RUNNING MUD OR HYDRAULIC ROTATION DRILLING RIG—Drilling rigs vary according to the depth of the well, although sometimes longer drill pipes are used to reach a greater depth with the same rig. Derricks are quickly set up, with the pieces assembled on the ground and the derrick then raised into position. The derrick's framework supports the entire drilling rig.

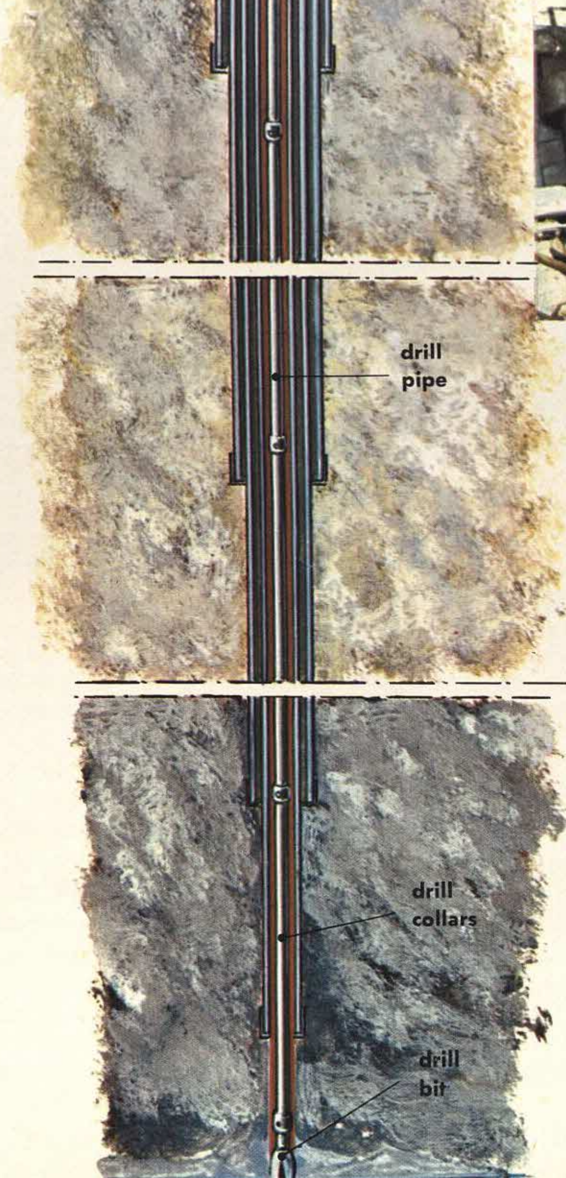
2

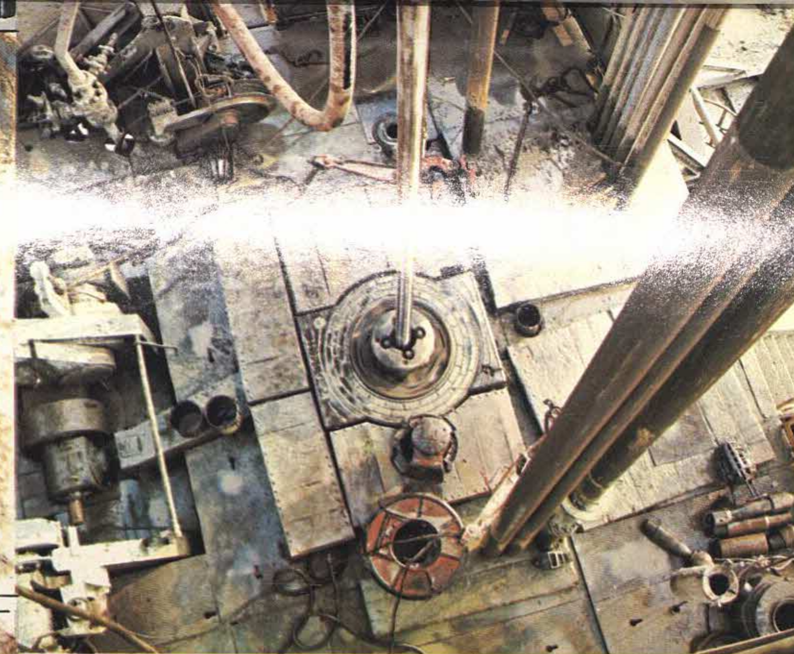
through the nozzles in the drill bit, from where it is pushed to the surface again along the outside of the stems.

On the surface, the mud flows over a shale shaker, a sort of vibrating sieve, and then into the collecting tanks. From there it is again collected and recirculated by the suction pumps.

Drilling mud has several uses: (1) it forms a kind of plaster on the sides of the well, preventing the sides from crumbling and blocking the shaft; (2) it prevents the violent eruption of methane, oil, or water as the drill bit bites through the rock formations—if fluid under pressure is found, the pressure of the mud is increased to create an equilibrium, thus preventing an eruption; (3) it lubricates and cools the bit, reducing the wear on it; (4) it carries the rock cuttings to the surface as they are formed.

Other safety devices, flow regulators, and controls are located near or under the derrick floor, while at the side are casings, various pieces of equipment, and measuring and testing instruments. The casings are laid out in a predetermined order to expedite their use.





THE COURSE OF THE MUD—The mud, passing through the tube shown in the illustration, is carried to the swivel or injection head, from which it is forced down through the drill pipes. The illustration also shows the turntable through which the system of drill pipes passes while it is rotating at full speed. 3

80 to 100 m (about 260 to 330 ft). The brush is cleared and access roads are checked to see if they are adequate for the passage of large vehicles. Provision is made for water supply, electrical power, and communications. Only after these preliminary procedures have been effected can the search for the petroleum begin.

The first step is the construction of a wooden platform, bolted to a foundation, sturdy enough to support a very large load, including the oil derrick.

The most common method of drilling today—the rotary drilling method—requires thousands of feet of drill pipe and well casing, a derrick, a power source or prime mover, drill bits, drilling muds and chemicals, and a team of several men.

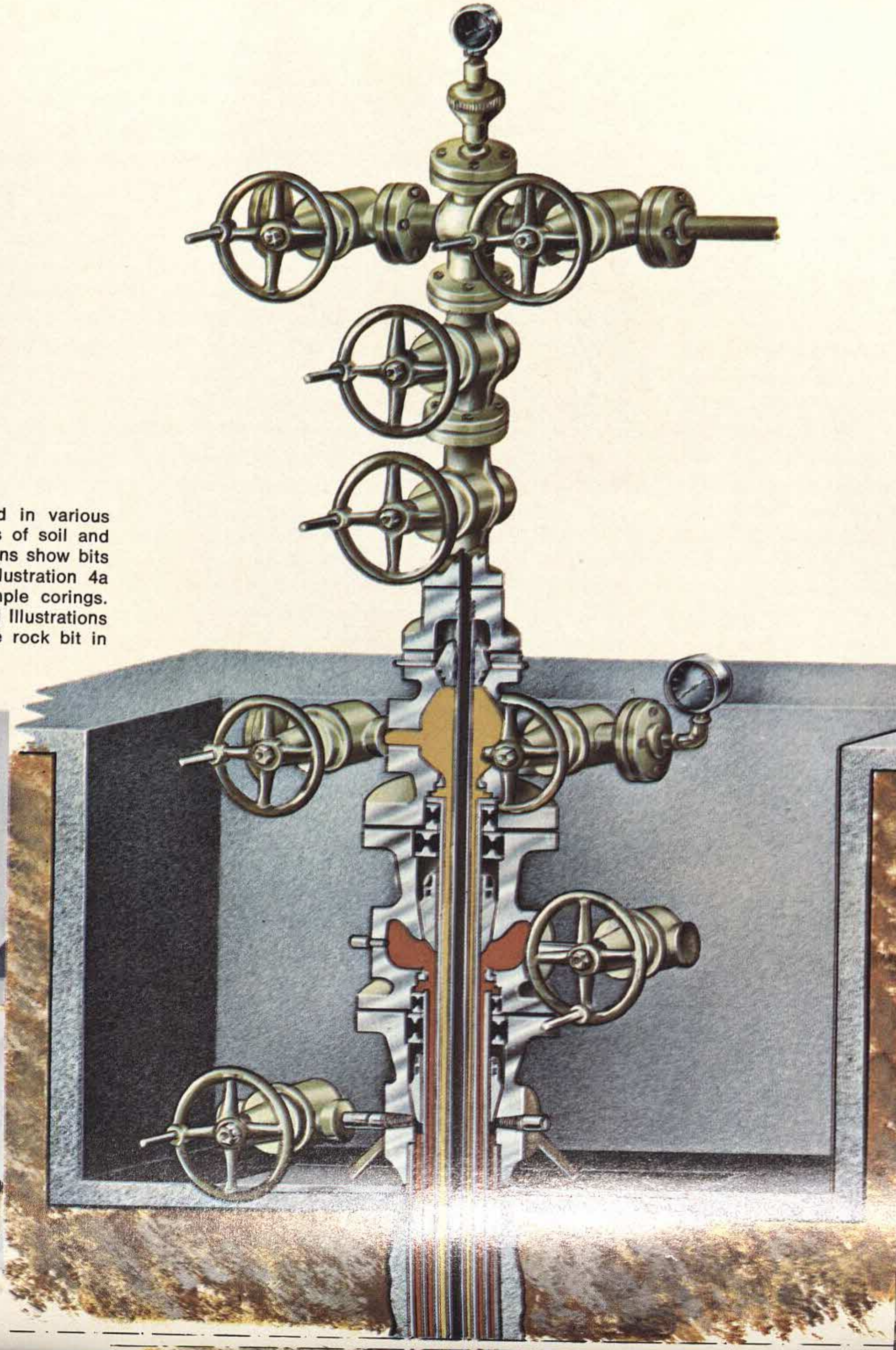
Wells drilled in the early 1960s in the United States descended on the average to a depth of more than 4,400 ft, at an average cost of $15 per ft; an average well thus involved an expenditure of about $65,000. In underdeveloped countries, where work often has to be done in desert, jungle, or swamp, and where the oil explorer has to provide all the utilities, the drilling of a single well may cost millions of dollars.

BITS—These tools are designed in various shapes for use in different kinds of soil and at different depths. The illustrations show bits made of case-hardened steel. Illustration 4a is a diamond bit for taking sample corings. Illustration 4b is a Zublin bit; and Illustrations 4c and 4d are the same tri-cone rock bit in different positions.

4a

4b

4c



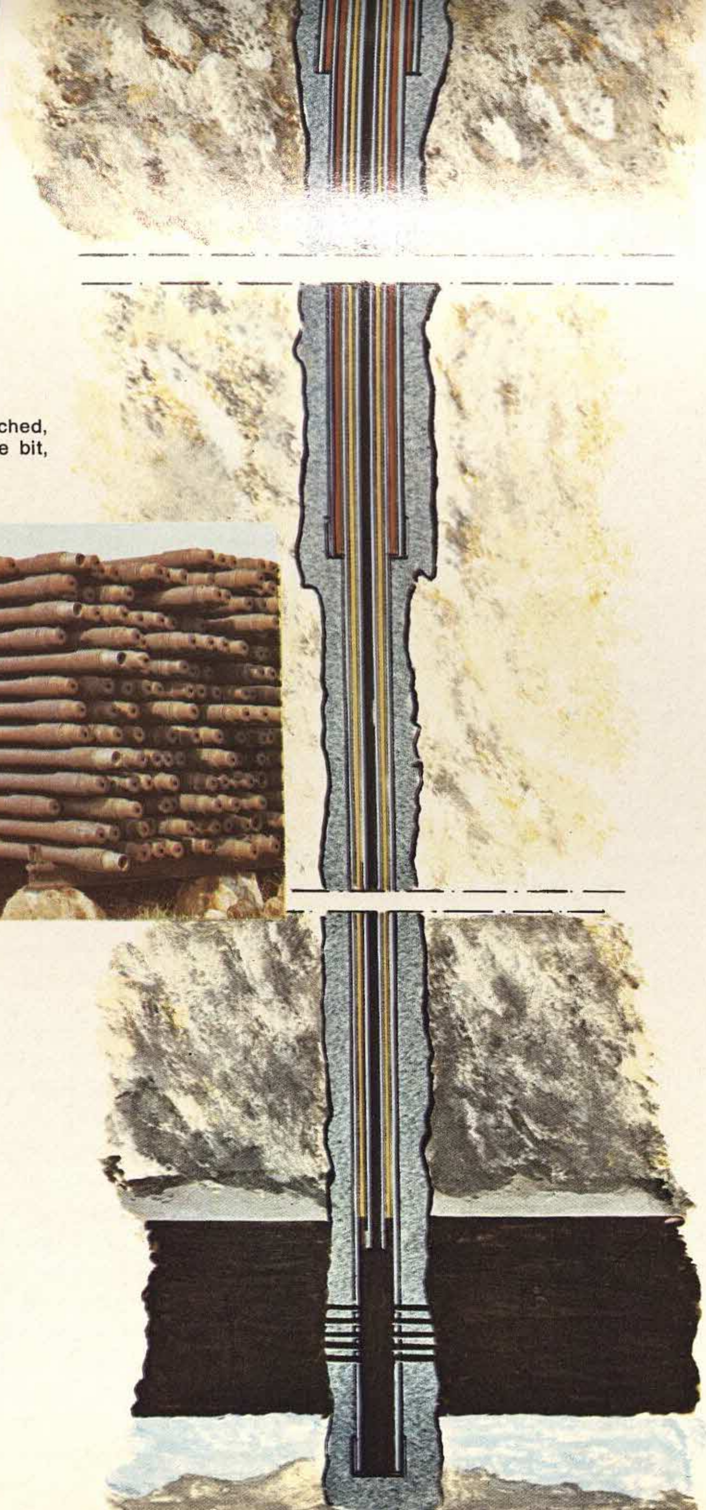
DRILLING—The three photographs show steps in the drilling process; the large diagram is a cross section of a producing oil well with the cruciform distribution head (often called the "Christmas tree") at ground level.

The first step in drilling is mounting a 44 to 45 cm (about 18 in.) bit directly on the kelly, which is supported by a strong cable and inserted through the turntable (Illustration 6a). In this way the well is drilled to a depth of about 10 m (about 33 ft), which is the length of the kelly. Then the kelly is lifted from the well and a drill pipe is attached between it and the bit (Illustration 6b). The well can then be drilled as deep as the combined length of kelly and drill pipe. Another drill pipe is then added.

When the well reaches a certain depth, the drilling bit and pipes are removed and the first section of casing is put into place. Cement is forced down through the casing and up around the sides to fill any channels through which fluids might escape. Drilling is continued with a smaller bit, and a casing of smaller diameter is inserted through the first. The last step (Illustration 6c) is introducing a production pipe about 7 cm (about 3 in.) in diameter; it extends from the distribution head at ground level down to the perforated concrete casing at the bottom of the well.

6a

4d



DRILL PIPES—These sections are attached, one at a time, between the kelly and the bit, until the desired depth is reached.

5

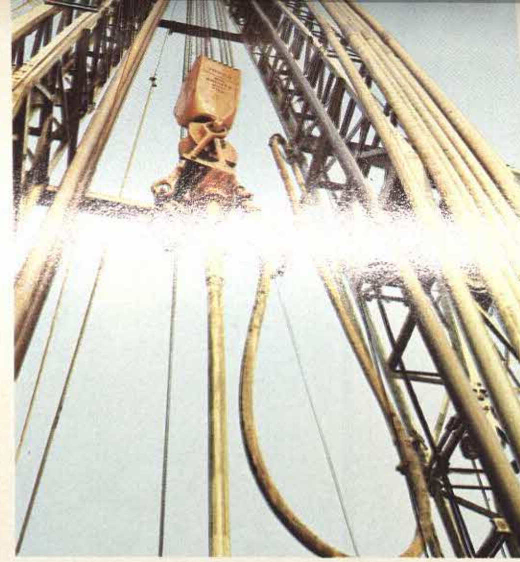
mud

gas

cement

oil

6b



6c

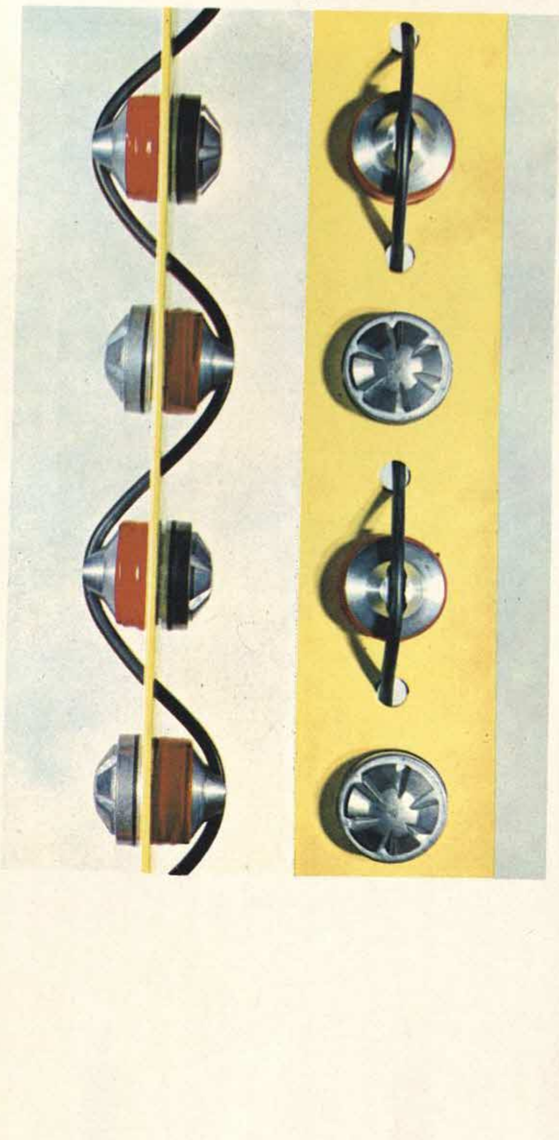
8

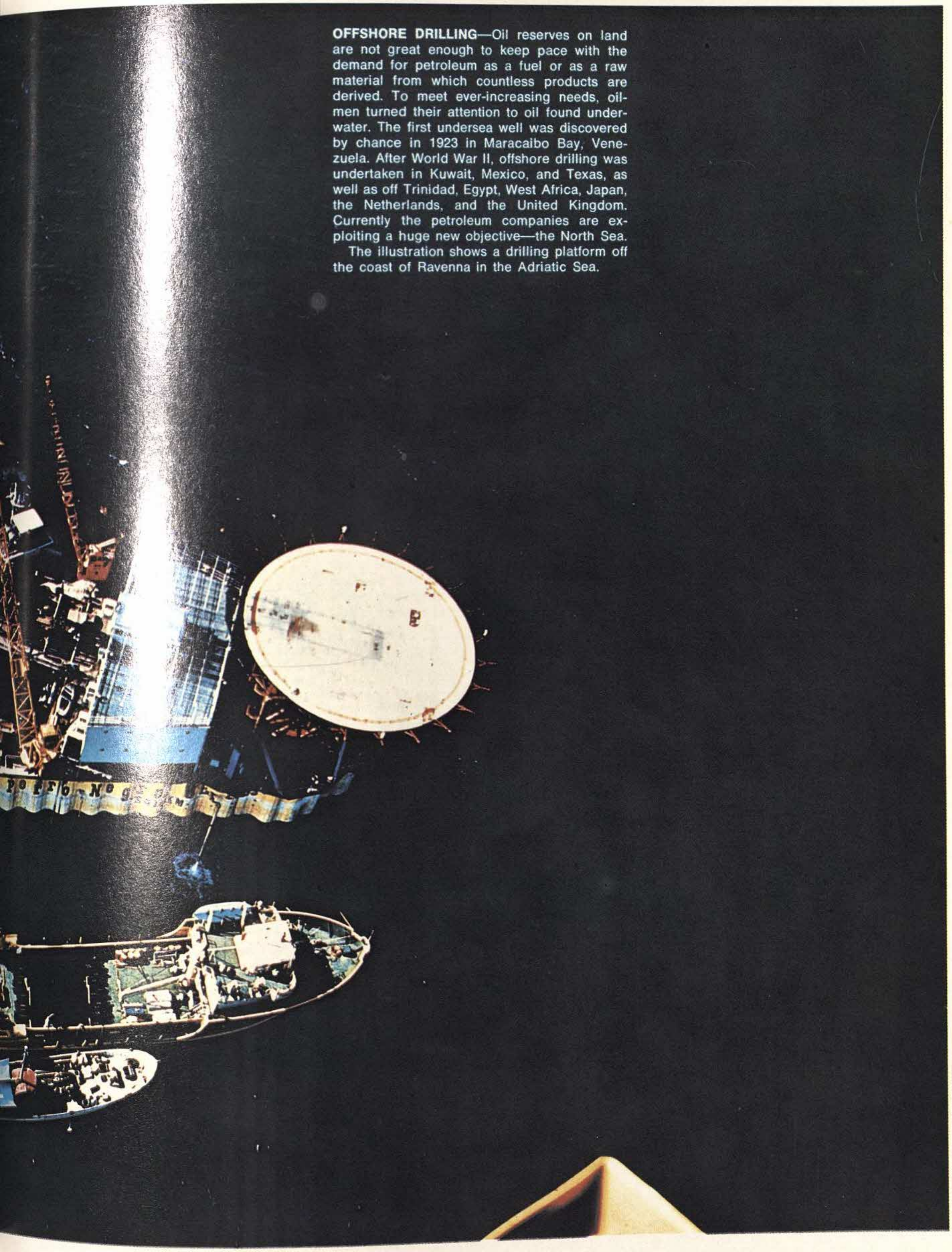
7

PERFORATION GUNS—After it has been established that there is oil or gas at the bottom of the well, the shaft is sunk still deeper into the petroleum-bearing stratum, and then the ending is cased with concrete. Two days pass while the concrete hardens.

Illustrations 7a and 7b are two views of a special gun that is lowered to the bottom of the well to perforate the concrete casing. Each unit of the gun contains a charge of explosive. When the gun is fired by electricity, each unit shoots a hole in the concrete casing. Through these holes, oil or gas flows into the well.

a b





OFFSHORE DRILLING—Oil reserves on land are not great enough to keep pace with the demand for petroleum as a fuel or as a raw material from which countless products are derived. To meet ever-increasing needs, oilmen turned their attention to oil found underwater. The first undersea well was discovered by chance in 1923 in Maracaibo Bay, Venezuela. After World War II, offshore drilling was undertaken in Kuwait, Mexico, and Texas, as well as off Trinidad, Egypt, West Africa, Japan, the Netherlands, and the United Kingdom. Currently the petroleum companies are exploiting a huge new objective—the North Sea.

The illustration shows a drilling platform off the coast of Ravenna in the Adriatic Sea.

PHOTOGRAPHY—I | developing photographic negatives

The typical commercial film-processing laboratory usually produces negatives of high quality at a modest cost and is quite convenient. An individual would be required to produce many rolls of film if he desires to develop his own film for the purpose of saving expenses. Most people who develop their own film do so for their own enjoyment and individual control over the developing process. The process is simple and requires very little equipment. The photographer who develops his own film is capable of making corrections for lighting conditions as the photograph was exposed. A negative of very high contrast can be developed in such a fashion that the contrast is lowered, or a low contrast negative can be developed to increase its contrast. The basic steps listed in the illustrations show the processing of a glass plate negative in darkroom trays. Sheet film and the more typical roll films are developed by the same basic steps. The person developing his own film should always pay close attention to the directions packaged with the film or developing chemicals. Almost all commercially available films can be developed with developers and fixing solutions that can be purchased as premeasured mixes. The experienced film developer may wish to experiment with solutions of his own de-

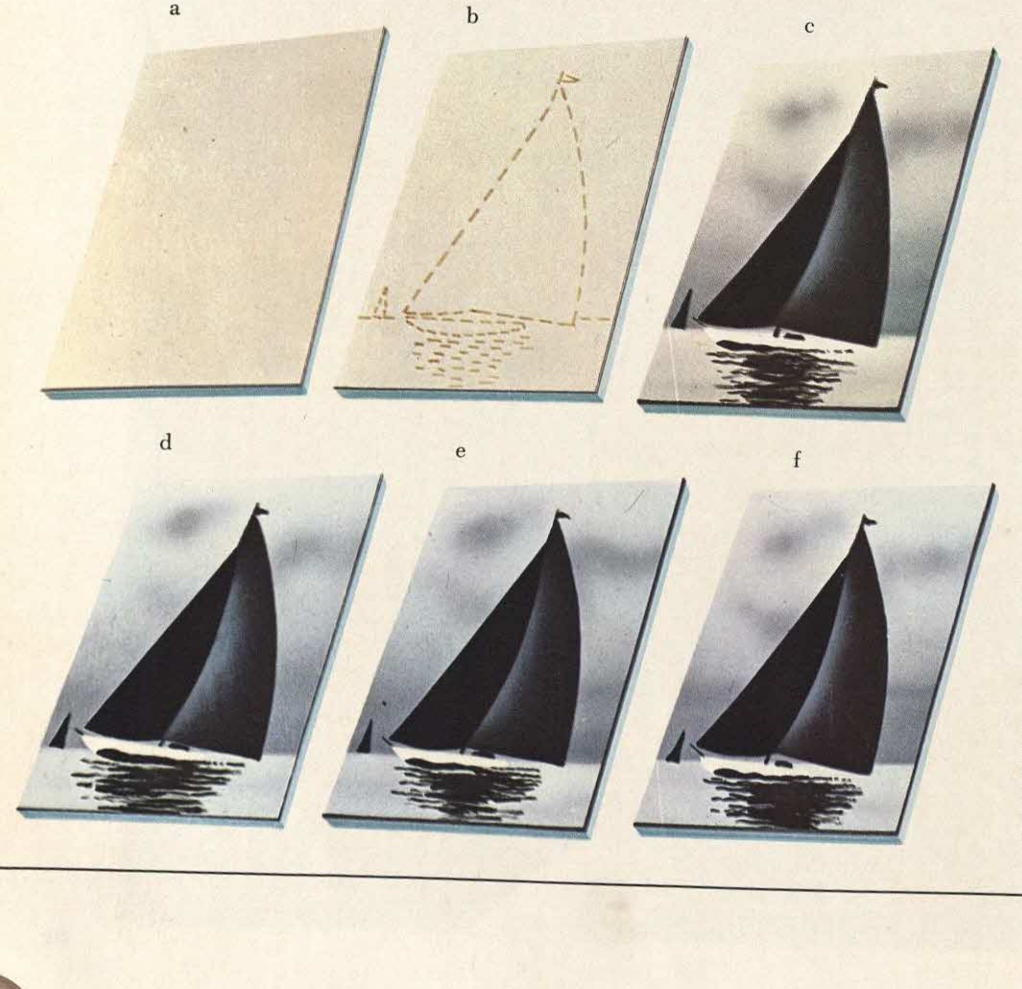
1

THE DEVELOPING PROCESS—The six illustrations indicate the six steps required to process a photographic film or plate. Though the illustration shows a sheet-film plate that is used in a press or view camera, the processes indicated are the same for any type of roll or sheet film. Visible changes occur only in the developing step (Illustration 1c) and the fixing step (1d).

The unexposed film or plate (Illustration 1a) is usually buff in color and contains all of the light-sensitive silver bromide. The silver bromide is suspended in a layer of gelatin that is uniformly coated on a suitable backing. Although early photographic plates were made of glass, today most film is bonded to some type of plastic or acetate backing. Many commercial films are offered for sale in a choice of backings. A thin acetate base is used for the common roll film as it must be flexible at all temperatures so that it will roll on camera spools.

There is no visible difference between an unexposed film and a film that has been exposed to light; however, subtle differences could be noted in the grains of silver bromide (Illustration 1b). The light has started a chemical process that will cause the silver in the silver bromide to be reduced to grains of metallic silver. This process will be continued in the developing stage so that the areas of the film exposed to light will be black in color and opaque to light. The film or plate that has been exposed to light is said to contain a latent image—an image that can be further developed chemically.

a b c d e f



If the film or plate is then placed in the developing solution, the latent image will become visible as small grains of metallic silver (Illustration 1c). The amount of light striking a particular area on the film determines the amount of metallic silver formed in the developing process. This effect produces a negative image—it is dark in areas where much light was present, and it is lighter in areas where little light was present. The overall developed film or plate remains buff in color, as the undeveloped silver bromide is present in the emulsion. This buff color will cause the entire film to be opaque to light.

The silver bromide emulsion coating on the film is sensitive to light until the fixing process is completed (Illustration 1d). Hypo or sodium thiosulfate is the prominent active ingredient in the fixing solution. The unreduced silver bromide in the emulsion is dissolved by the chemicals in the fixing solution. After the fixing process is completed, areas of the film that have no silver deposits become transparent, and areas that have high silver content will be completely opaque. Most areas of common negatives have a density that lies between completely transparent and completely opaque.

The fixing solution often contains a hardening agent that will cause the gelatin emulsion to become more rigid and, hence, more durable. Other chemicals are often added to the fixing solution to give desired effects. The film or plate is no longer sensitive to light after it has been placed in the fixing solution.

After completing the fixing process, the emulsion is free of the silver bromide; however, the chemicals of the fixing solution remain in the emulsion. A thorough washing is required to remove these chemicals (Illustration 1e). If the chemicals are allowed to remain in the emulsion, they will crystallize on drying and cause the negative to have a mottled effect. A cold running water wash is used for about fifteen minutes to free the emulsion of any residual chemicals. Warm water would cause the delicate emulsion to crack and separate from the backing.

After the washing, the negative still contains water and is quite susceptible to scratching. The emulsion is swollen and soft and must be thoroughly dried before handling. The processed negative should be dried in air that is free from lint and dust as they will readily adhere to the softened emulsion. The dried negative (Illustration 1f) may now be stored or used to make a positive picture or plate.

2

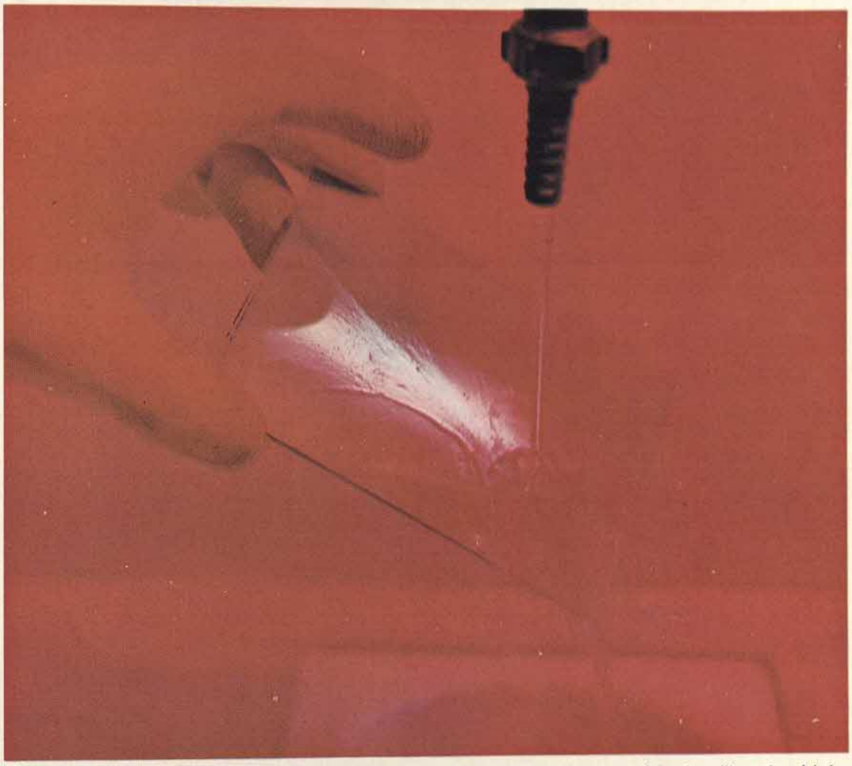
sign. The more a photographer delves into the developing process the more he enjoys the production of high-quality photographs.

PREPARATIONS FOR DEVELOPING NEGATIVES

Photographic negatives should be developed in a completely dark room. If the film is orthochromatic, a safelight may be used during the developing process. Most orthochromatic films are sensitive to blue light only, and a safelight, which is red or amber, may be used. Many of our films today are panchromatic and are sensitive to light over the entire spectrum–they should be developed in total darkness.

A developing solution and a fixing solution should be selected from the list of suggested developing chemicals packaged with the film. After the chemicals are prepared, they should be placed in trays and their temperature checked. Developers are designed to be most effective at a temperature of 20° C (68° F), but it is usually safe to use chemicals that are somewhere between 18° and 24° C (65° and 75° F). Only three trays are required to hold the three solutions for developing the film. Most people who develop film place developer in the left, stop bath in the center, and the fixing solution in the right tray. Forceps or tongs should be used to transfer the film from tray to tray during the developing process. Since developing takes place in the dark, it is almost necessary to touch the film during the process. Care should be taken not to drip chemicals from the fixer tray into the developer tray as this can ruin the film.

The development of roll film should be accomplished in a special tank designed for that specific purpose. These tanks are designed to hold the film in such a way that it does not touch the surface of the tank or adjacent portions of the film. Once closed, these tanks are lighttight and the developing takes place in room light. The chemicals used to process the film are alternately poured into and out of the tank for the proper times. The developed roll of film is then dried by hanging in a dust-free place. Negatives are usually dried in moving, warm air.



WETTING THE NEGATIVE—The negative should be wet before it is placed in the developing solutions. A thoroughly wet negative will accept the developing solutions uniformly and will not contain areas that develop for longer periods of time than others. The illustration shows a red light being used as a safelight. The directions with the film should be followed in choosing a safelight. Most modern panchromatic films require that the entire developing process be carried out in total darkness. Often a very dim safelight may be used for a few seconds during the development to facilitate inspection of the negative.

3

THE DEVELOPING PROCESS—The selection of the developing solution again depends on the directions packaged with the film. The developer chosen determines the tone of the negative. Some developers will give cold tones—crisp sharp whites and blacks, while other developers are chosen to give warm tones—a mellowing of off-white to off-black. Warm-tone developers are most often chosen for portraits; cold-tone developers lend themselves to sport event photography. If the eventual print is to be used in a publication, a cold-tone developer is often recommended.

The negative should be placed in the tray with the emulsion side up to keep it from being scratched on the bottom of the tray. Care should be taken to handle the negative only by the edges to prevent scratching the surface or partial warming in certain areas. The developing process is quite temperature-sensitive, and areas of the film that are warmer will develop to a greater extent. Gentle agitation is required during the entire development to ensure that fresh chemical is always in contact with the film. This agitation can be accomplished by a slow rocking of the developing tray. Timing the developing process is quite critical, and the directions packaged with the film should be followed closely. It is best to use a clock with a preset alarm. Ordinary clocks cannot be read in total darkness, and those with phosphorescent numerals could give off light that would cause a fogging of the film.

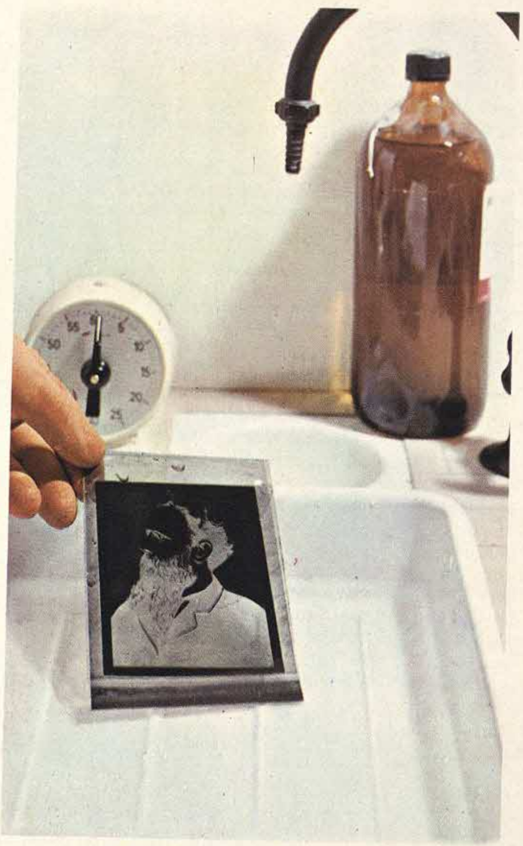
4

THE STOP BATH—The developed film should be removed from the developer at the indicated time and then placed in the fixing solution. A stop bath is often used between the developing and fixing processes. The purpose of the stop bath is twofold: the developing process is abruptly halted in the stop bath, and the life of the fixing solution is extended. The developing solution is basic or alkaline while the fixing solution is slightly acidic. The stop bath is an acidic solution that abruptly neutralizes the alkaline developing solution; this causes the development to stop. If no stop bath is used, the development and neutralization would occur in the fixing solution—and this would rapidly deplete the strength of the fixing solution.

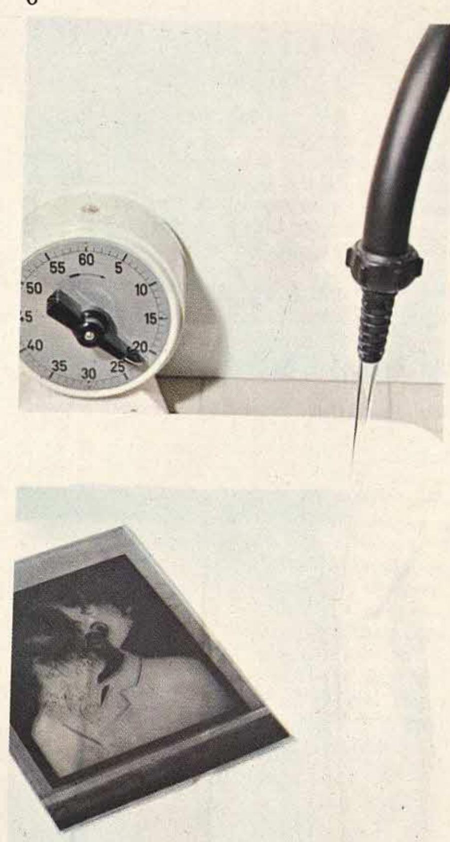
DRYING—The film should be air-dried and protected from lint and dust. The illustration shows a glass plate leaning emulsion side down against a corner. Nonrigid films are most often dried by hanging them in a place free of drafts and dust. A fan is not recommended to hurry the drying process because rapidly moving air holds much more dust and lint than calm air. Heating is also not recommended as the gelatin emulsion will separate from the backing and wrinkle.

7

5



6



THE FIXING SOLUTION—This bath dissolves any residual silver bromide. Although silver bromide is not soluble in water, it will readily dissolve in the solution of sodium thiosulfate. After the developed negative has been placed in the fixing solution and the process of dissolving the excess silver bromide has been completed, the film may be exposed to normal light. The fixing process requires about fifteen minutes to complete, but the film may be inspected after the first five minutes. Fixing, like developing, is temperature-dependent, and the solution should be maintained at a constant temperature. Most developers and fixing solutions are recommended for use at 65° to 75° F.

WASHING—After the steps of developing and fixing are completed, the negative should be washed for at least one half hour in cool running water. The running water removes the chemicals left in the emulsion by the previous steps. Sheet or plates of film are effectively washed in a tray, as in the illustration. Care should be taken to prevent a strong stream of water from striking the emulsion directly as the emulsion is still soft and may be easily damaged.

8

STORING — After drying, the film can be handled without excessive precautions. It remains a delicate object and, hence, it is best to take precautions when storing. Plastic or wax paper storage envelopes are most often used for storing negatives. Ordinary envelopes are sources of lint, and the negative should be as lint-free as possible if it is to be used to produce a satisfactory photograph.

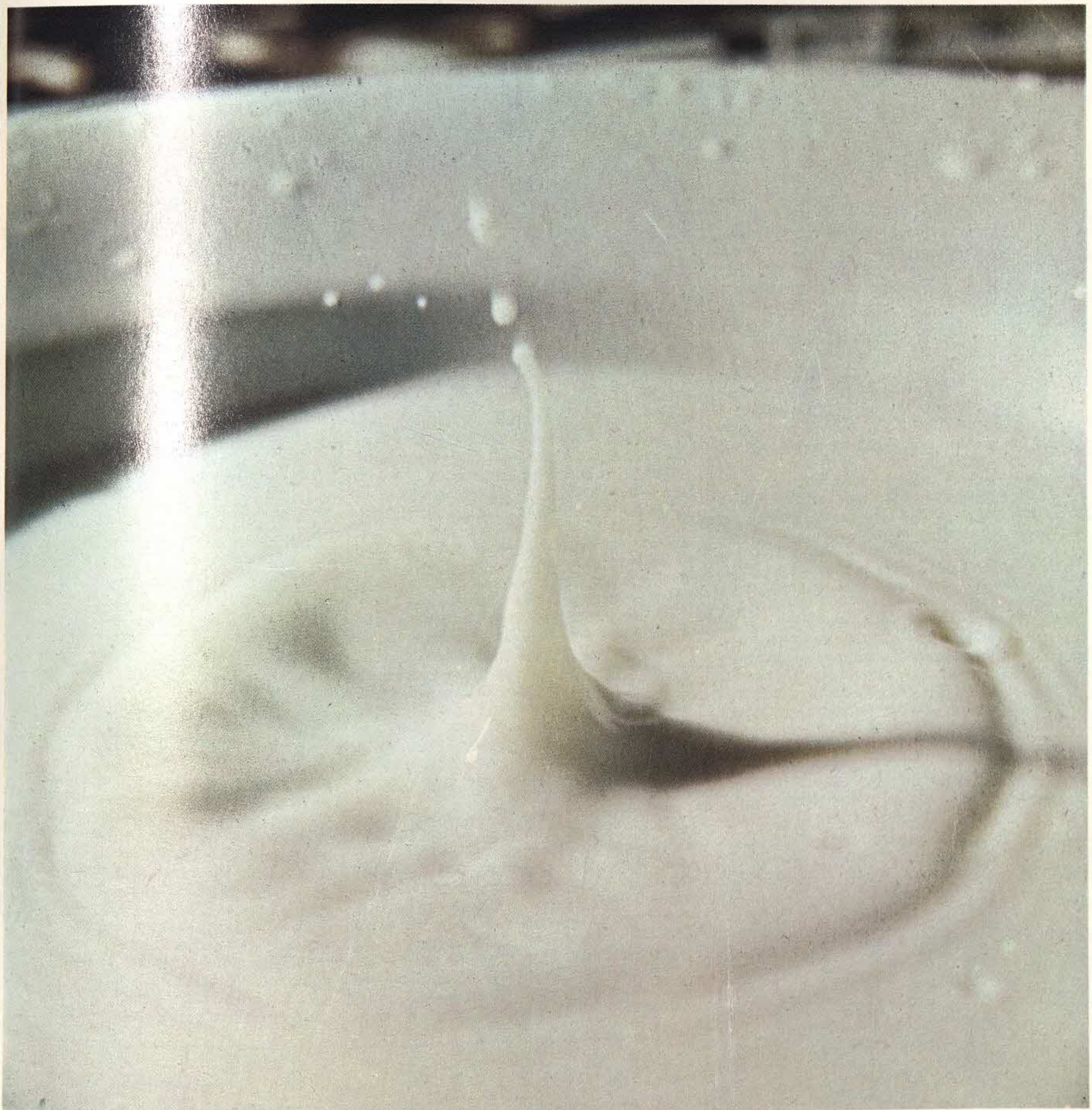
PHOTOGRAPHY—II | high-speed techniques

The stroboscope is an instrument that is often used in observing phenomena that are either fast-moving or of short duration. The stroboscope consists of a flash of very bright light that is of very short duration. The methods illustrated here are often used to record ultra-short time interval images. The typical electronic flash used by many photographers has a flash duration of about one thousandth of a second. The time lapses for ultra-high-speed photography are in the range of one ten-thousandth to one hundred-thousandth of a second. These short time intervals are required to photograph a moving bullet or a bursting balloon. The typical electronic flash is much too long for either of these phenomena. A speeding rifle bullet would travel nearly three feet in a thousandth of a second and would be recorded as a blur by the conventional electronic flash. The Ruhmkorff induction coil is capable of giving off light of low intensity but at very short time intervals—such a light is suitable for ultra-high-speed photography.

Certain advances in electronics have resulted in more efficient electroluminescent devices that operate in the visible region and that might have useful application in high-speed photography. Laboratories in many parts of the world have been investigating electroluminescent di-

1

a



odes as sources for visible light. Three-part compounds, such as those containing gallium, arsenic, and phosphorus, or gallium, aluminum, and phosphorus, have been used to produce red emitters; silicon carbide has been used to generate blue light. In addition to their use as point sources of light that require only a few milliwatts of electric power, diodes have a high-speed response.

THE ELECTRONIC FLASH — An electronic flash is a convenient device for giving brief and intense flashes of light for stroboscopic photography. The flashes are produced by an electrical discharge passing through a xenon-filled tube. The impulse of the electric current produces a flash of 1/500 to 1/2,000 of a second. The magnesium-filled flash bulb gives off much illumination but has a duration of about 1/25 second—this is much too slow to "freeze" rapidly moving objects. The electronic flash can be used to photograph events that occur too rapidly to be seen directly.

Illustration 1a shows a drop of milk being added to a glass of milk. If the drop falls from a sufficient height, the spout formed in the center can be observed. It is impossible to observe much about this spout or the surface of the milk around the spout without the sharp picture produced by an electronic flash.

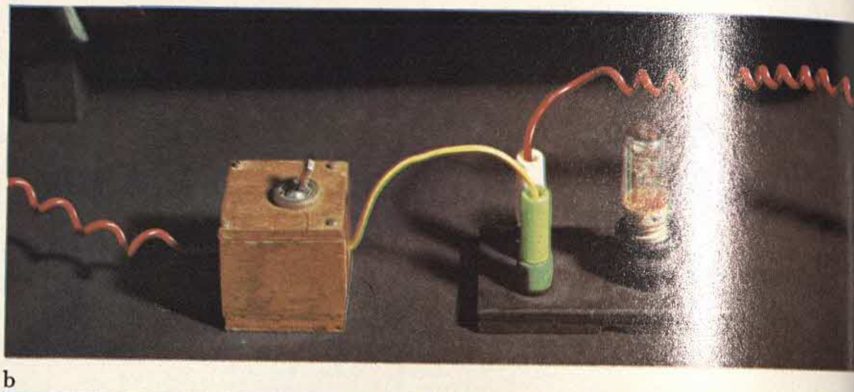
Illustration 1b shows a light bulb being shattered. When an object is struck, it either resists the blow or is shattered. The shattered pieces fly toward the inside of the bulb at a speed so great that they cannot be followed by the unaided eye. It can be noted with the flash photograph that the broken pieces are indeed moving toward the center.

Photographs of this type require special camera settings. The camera is usually placed in an unlighted room with the shutter open. The xenon-filled flash tube and the camera are both pointed at the subject to be photographed. The flash was activated in this illustration by allowing the hammer to make contact with the bulb. The wires shown in the illustration are connected to the shutter contact on the electronic flash, and as contact is first made, the flash illuminates the hammer and the bulb simultaneously. This procedure is quite simple and can be accomplished with very little difficulty. It is possible to set the flash to go off at different times and even to repeat at fixed intervals. This allows the observer to record the movement of the shattered pieces of the bulb at intervals.

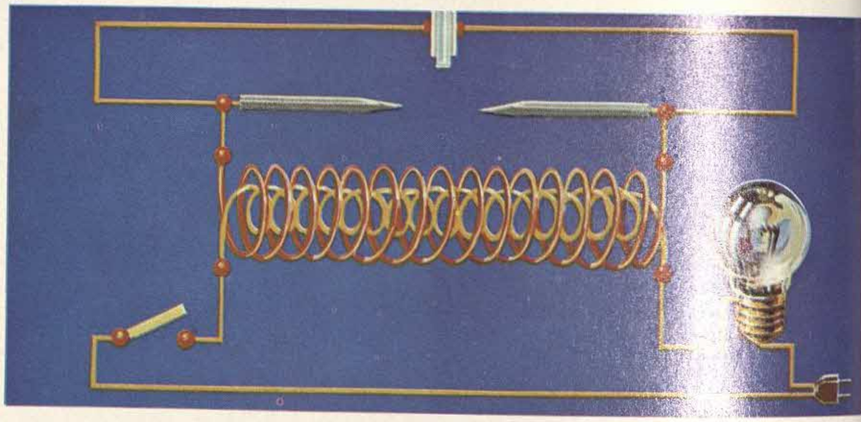
b

2

a



b



c

ILLUMINATION BY SPARKS—A flashing electric spark is of shorter duration than the flash of a xenon-filled flashtube. It can be produced by making a simple modification of a Ruhmkorff induction coil. Its 50,000-volt flash is rather dim and gives off a weak violet light. The flashes from the induction coil occur at irregular intervals and continuously, depending on the particular makeup of the electric vibrator. For photographic purposes, the induction coil must undergo some modifications.

The vibrator of the Ruhmkorff induction coil is eliminated and replaced by a push button switch and a light bulb (Illustration 2a). The switch is wired so that the primary circuit of the coil is closed until the switch is depressed. The light bulb in the circuit acts as a load resistor. This load resistor is required to protect the induction coil primary circuit, as in normal operation it is only energized with the periodicity of the vibrator and then only for very short periods of time.

Illustration 2b shows the complete wiring for the modified induction coil circuit. The vibrator has been replaced with a switch and a load resistance of a light bulb. The top of the illustration shows the parallel wiring of a capacitor, which is added to strengthen the flash. The capacitor is wired directly to the spark gap as is the output of the induction coil secondary circuit.

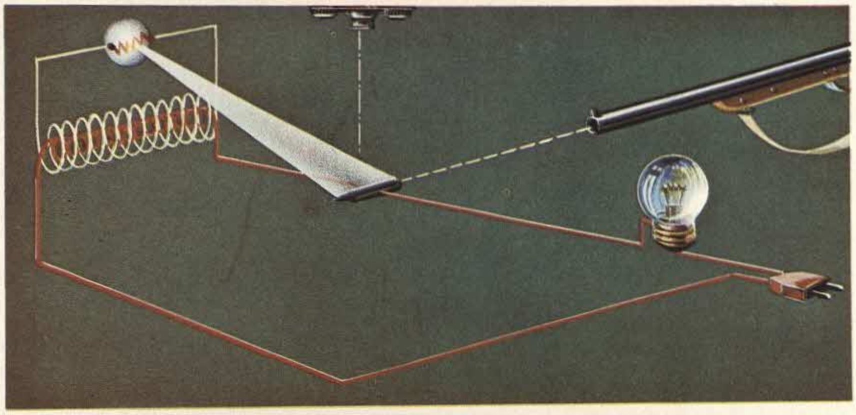
The capacitor (Illustration 2c) is one sheet of glass about one foot square. Aluminum foil is placed on the surfaces of the glass sheet, and care is taken that it does not extend to the edges. Because the glass must withstand a voltage of 50,000 volts, it should be about ¼ inch thick. The edges of the capacitor are coated with paraffin. The completed capacitor is placed in wood blocks so that it stands in a vertical position. These blocks are also coated with paraffin. Electrical connections are made to the aluminum foil sheets on either side of the capacitor.

Extreme care should be taken when energizing the induction coil circuit. The spark provided is of short duration and of very low current, but the voltage can be fatal.

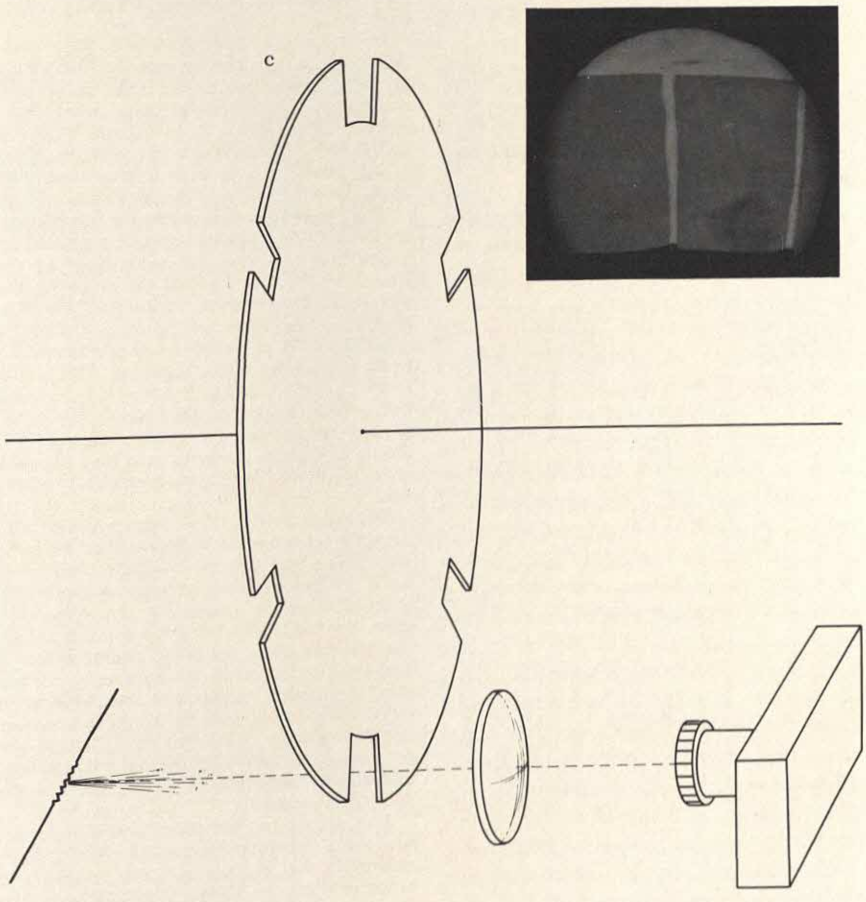
3

a

b



c



OBSERVING ULTRA RAPID PHENOMENA— By using spark flashes it is possible to observe and record phenomena that occur at intervals up to one millionth of a second. By using the modified Ruhmkorff induction coil and a darkened room, many short interval photographs are possible.

Illustration 3a shows the light reflected from the wings of a dragonfly. The dragonfly is trapped and caged, flapping its wings in a violent attempt to escape. The spark from the induction coil can easily be used to "freeze" this motion.

Illustration 3b shows the setup required to observe and photograph a bullet in flight. An electrical connection in the primary circuit of the induction coil is broken by the bullet—in effect, the bullet will take its own picture. Fast-moving and nonrepetitive events must be recorded in this fashion. It would be nearly impossible for the photographer to push the button at the exact interval that the bullet is in front of the camera lens.

Illustration 3c shows a method that can be used to measure the duration of the flash from the induction coil. A notched disk is mounted between the spark gap and the lens of the camera. Two photographs are taken of the disk using the discharge from the induction coil. The first photograph is taken with the disk stationary and the second with the disk moving at a known speed. The speed should be as high as possible and at least 3,000 rpm. The notches in the photograph of the moving disk will be larger than those of the stopped disk. A comparison of the two photographs will be needed along with the speed of the disk to calculate the exact duration of the flash. The photograph in the illustration shows notches in a disk that is rotating at 3,000 rpm and the flash duration was calculated to be 5 microseconds.

PHOTOGRAPHY—III | processes of color photography

Methods for producing color photographs were developed shortly after the invention of black and white photography. Almost as soon as black and white photography became a reality, many photographers attempted to produce color emulsions. Nearly all of the patents for color photography were granted between 1850 and 1900. Nevertheless, none of the emulsions and processes could be used to meet practical demands for a method of color photography because production techniques were too primitive. Only after 1930 did the practical use of color film become a reality; however, the technology required to produce color films in quantity required time. It was also necessary to resolve the problems of commercial production, developing, and printing of color films.

Color photography can be separated into two basic types. The color separation method involves the use of black and white panchromatic films to produce negatives with colored filters. The second method uses a combination of photographic emulsions sandwiched into one negative. The typical color films of today are of the latter type.

The color separation method involves three separate exposures with proper color filters. The negatives may be produced by successive exposures with the same camera, or they may be simultaneous exposures with different cameras. In either case the three negatives produced are used to reconstruct the color of image. Such an image can be projected with proper color filters, or it may be used to produce typographic plates for color printing. Most color illustrations in this book are produced by this method. Each illustration is printed four times with plates formed from color separation negatives (the additional plate is black—it is added to complete the illustration). Each plate is inked with the color corresponding to the original separation negative.

The color photograph or transparency most widely used today in both still photography and motion pictures is the sandwiched emulsion or the tripack. The photograph may be formed by one of two methods. The tripack can produce a color negative that is used to print a color photograph, or a color reversal process may be used that converts the film into a transparency or color positive.

The following illustrations explain the methods used in color separation photography, and the methods used to reconstruct images from separation negatives.

TRICOLOR SEPARATION—A color image may be produced by combining three black and white images that have been photographed with special colored filters. This method of color photography is widely used today to produce printed images; however, it is seldom used to produce projected images.

In the illustration, the object **a** is a flower that has been chosen for its particular colors. The flower and its pot contain the primary colors red, blue, and yellow; the green secondary color is the additive color of yellow and blue. The example does not contain all the color ranges of the visible spectrum, but the process of combining the primary colors can be extended to include the entire spectrum.

Three cameras are fitted with filters that are used to produce images corresponding to the primary colors. The filter colors are chosen so that their total color, when combined, is white. Filter **e** on camera **b** is magenta, filter **f** on camera **c** is green, and filter **g** on camera **d** is blue. A frosted glass has been placed in the focal plane of each camera so that the image produced by the camera can be viewed. The white background produces an image that is exactly the same color as the filter. The different parts of the object in the picture appear either as the color of the filter or black. The image on the frosted glass of camera **b** is black in the areas corresponding to the stem, leaves, and flower pot. This is due to the absence of magenta coloring in the blue and green portions of the picture. The background, petals, and the yellow center of the flower all contain magenta, and they produce an image that is exactly the same color as the filter.

The image produced on a panchromatic film is shown as **h** in the illustration; all areas of the object that contain magenta produce a black image, while those areas without magenta produce the transparent part of the negative. The negatives **i** and **l** are produced by the cameras using the green and blue filters. Negatives produced by this method are called color separation negatives. Each negative in this illustration has both black and clear areas, but it should be noted that most objects are not as pure in color as those in the illustration, and that the images most objects produce are of varying tones.

A closer look at the object and the respective color separation negatives reveals characteristics that are useful in the printing and projecting of the image. The colors of the object that correspond to exact filter colors produce negatives that are clear with the other filters. For example, the petals of the flower are pure magenta and the petals produce a clear image on the green and blue separation negatives. The yellow center of the flower is not a filter color, and the separation negative that it produces is clear only with the blue filter. An object of continual tone and color intensity produces a very complex set of color separation negatives.

If color separation negatives are to be used for projected tricolor images, they must be reversed. This process involves a simple printing of the negative to produce a positive of the original. The illustration shows these positive separations as **m, n,** and **o**. An image of proper tone and color intensity can be produced from these color separation positives by projecting each positive through the same color filter through which it was taken.

The image produced by each separation positive is identical to the image produced on the frosted glass at the top of the illustration. The positive **m** produces a projected image similar to the frosted glass b. If all three color separation positives are projected through their respective filters in proper alignment, the result is a projected image similar to the original object. This method of reconstruction of the colored image for projection, although popular for several decades, has been replaced by other more practical methods.

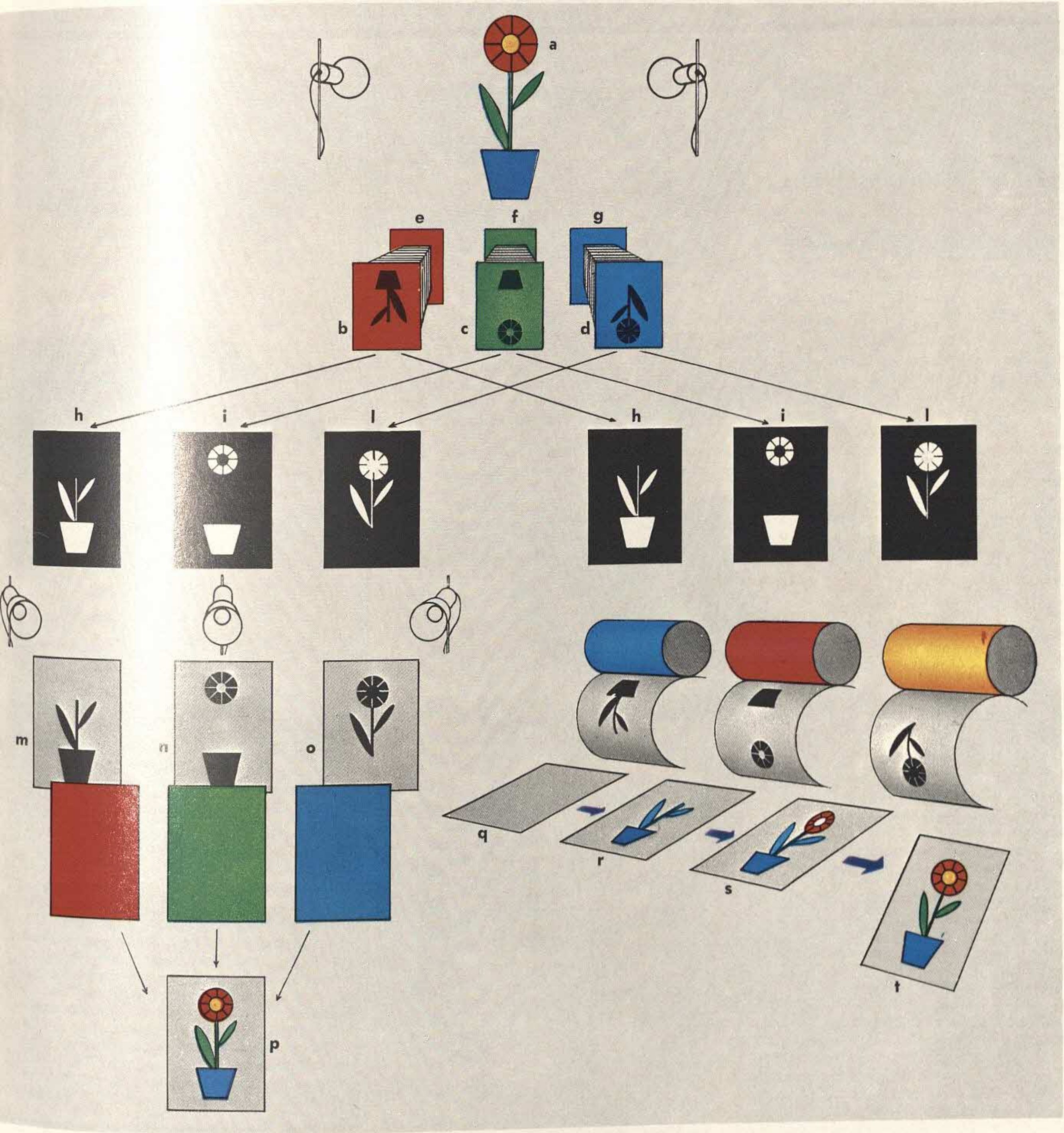
Printed images may be produced by tricolor separation using the same negatives **h, i,** and **l** as were used in producing the projected image. The three color separation negatives are used to expose and produce typographic plates. These plates, usually of zinc, are photographically exposed through the negatives. The zinc is etched away in areas that were not exposed to the light.

Typographic plates produced by this method are then coated with ink that is of a complementary color to the filter. Three printings are required to produce the colored image. Each typographic positive inked with its complementary color must print in turn on a white sheet of paper to produce the proper tonal renditions of the printed image. This method of color printing is widely used when a large quantity of photographic reproductions is required.

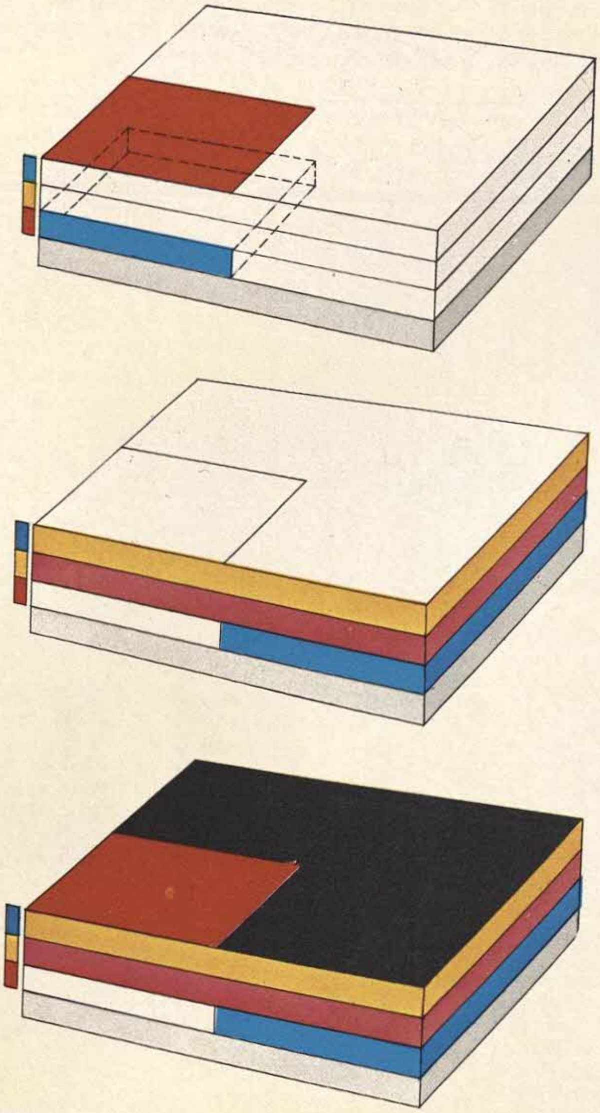
In practice it is important to have perfect standardization of the filter colors used for separation, of the gelatins used for projecting images, and of the inks used in the printing. A color produced by the overlapping of three basic colors is quite sensitive to the amount of each color present. Small irregularities in the developing of the films, in the alignment of the images, or in the inking can produce inferior results.

A detailed discussion of the development of the integral tripack explains the process used for color reversal of the sandwiched film. A number of manufacturers make films and printing materials of the types mentioned in a variety of forms including motion-picture films, sheet and roll films, cartridges, and cassettes.

1



2



THE INTEGRAL TRIPACK — The integral tripack is the film used for conventional color transparency positives for projection. Most color slide film utilizes this or similar methods to produce the photographic transparency. This illustration shows the makeup of such a tripack.

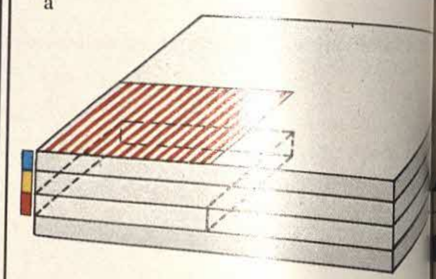
The tripack consists of three layers of light-sensitive photographic emulsion placed on a support or backing layer. The top emulsion layer is sensitive to blue light, the center layer to yellow light, and the lower layer to red light. The small insets to the left of the tripack indicate the color of light to which each layer is sensitive. These layers have the property of exposing the silver bromide only when light of the proper color is present. They also have the property of accepting the proper coloring agents in a precise quantity to produce good-quality photographic images. The amount of coloring accepted by the layer is dependent on the amount of silver bromide exposed in the photographic process.

In Illustration 2a, a red light has been projected on the lower left side of the film. This red light passes through the upper two layers of the film and impresses itself on the lower layer of the tripack. The processing involves developing of all layers of the film and the subsequent dissolving of the developed silver bromide. A portion of the lower layer is completely developed while all other areas of the layers undergo no change in the first developing process. After the film is exposed to white light, all of the previously unexposed areas become exposed. Those areas are developed and colored with dyes as shown in Illustration 2b; each layer is colored with its complementary dye. For example, the blue sensitive layer that was not exposed initially is now colored with its complementary color, yellow. The exposed portion of the red layer is now colorless because all of the silver bromide was dissolved before the initiation of the coloring process.

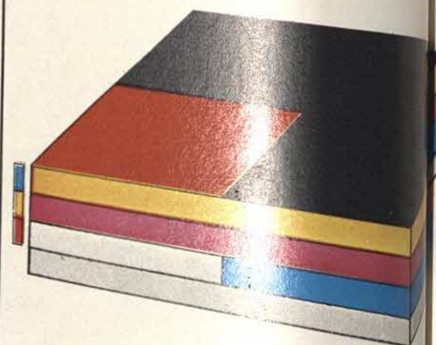
Illustration 2c shows the result of passing a white light through the film from bottom to top. The complementary colors of each layer absorb an amount of light that is proportional to their color and degree of exposure. The result is that all of the film appears black except in the area that was exposed to red light.

Color negative film is produced by the same type of color tripack. The color reversal is not produced in the film, but in the printing process.

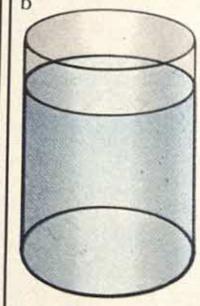
a



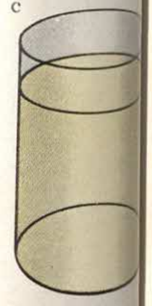
n



b



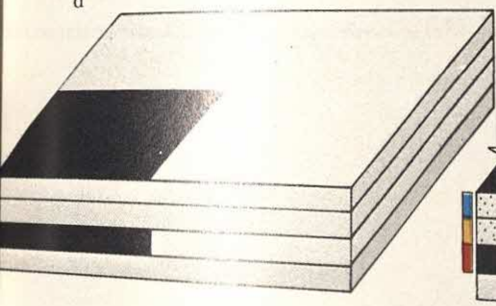
c



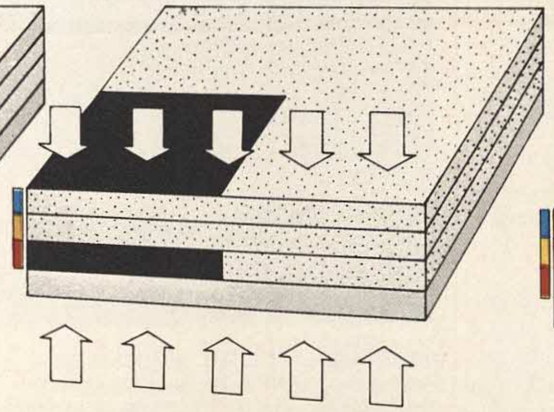
DEVELOPING A COLOR TRANSPARENCY — This illustration indicates the steps required to develop the film described in Illustration 2. The fragment of film has been exposed to red light in the lower left corner, and the light has completely saturated this one area. The example is simple; however, the process can be expanded to include complex colors and different shadings of those colors.

Illustration 3a shows the outline of the area that has been exposed to red light. The lower, red-sensitive layer contains a latent image of this exposure. The image does not become

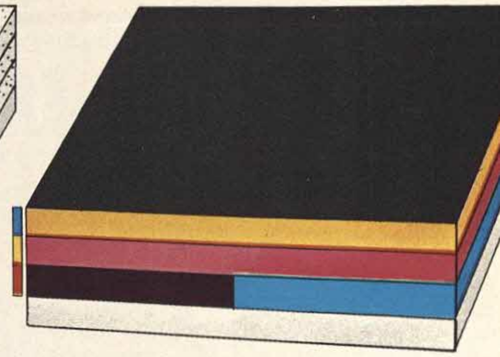
d



f



i



evident until the developing process has been completed.

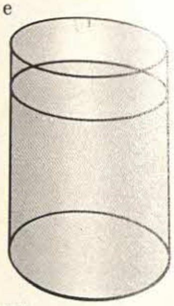
Illustration 3b shows the preconditioning bath, which serves to uniformly wet the film and impede the formation of bubbles in the developing process. Illustration 3c is the first developing bath.

Illustration 3d shows the film after it has been developed in the first developing bath. The layer of film that was exposed to red light is developed and the corresponding layer of the negative becomes black with deposits of silver. This example is of a layer that was completely saturated with red light only. The darkening would not have been complete had a typical photograph been used as the example.

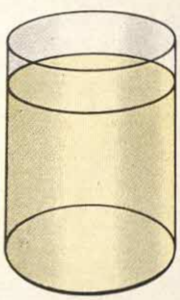
Illustration 3e shows the bath used to stop the developing process and simultaneously harden the film. (The developing process often weakens the gelatin layers of the emulsion.) All of the developing processes must be carried out at precise temperature and with precise timing, as the color quality of the image is quite sensitive. (The processes for the development of color films are much more critical than the corresponding processes for black and white film.)

Illustration 3f shows the film being re-exposed to white light. The light is allowed to enter the film from both the top and bottom. All of the layers not previously exposed and developed are now exposed to light in such a quantity that the layers become saturated. The example shows light causing a latent image to form in the upper two layers and in the

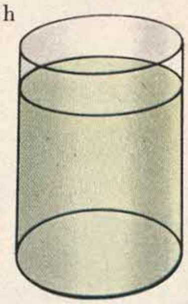
e



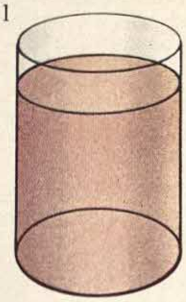
g



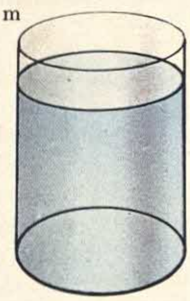
h



l



m



undeveloped portion of the lower or red layer.

Illustration 3g indicates the developing and coloring bath. As the latent image just formed during the re-exposure is developed, the color is brought out in the layer. The amount of color formed within a layer depends on the amount of silver being developed. The red layer that had been previously exposed and developed receives no new color. Only as the silver is being developed is the color reaction catalyzed.

Illustration 3i shows the film after this developing and coloring process. The entire film now appears black as all of the silver has been developed. This silver is now in the three emulsion layers, which also contain the colors that were formed during the second developing process.

Illustration 3l is the bath that dissolves the silver formed in each of the developing processes. No silver remains after the redissolving process; only the coloring dyes remain in the three layers. The film is then placed in a bath (Illustration 3m), which contains chemicals to harden and stabilize the emulsions.

Illustration 3n is the washed and dried transparency. The colors formed in the three layers that were not exposed to the red light combine to result in black, while the colored areas in the upper two layers form the red image. If white light had been included in the initial exposure, no color would have been formed in any of the layers, and the film would have been clear. The colors are chosen in such a manner that combinations of the three basic layers can be combined to reproduce images of objects having the entire visible spectrum of colors.

RADAR—I | basic principles

The term *radar* is of English origin; it is an acronym for "*ra*dio *d*etecting *a*nd *r*anging." Radar, therefore, is a means of detection and distance measurement using radio signals. The operating principle of radar is quite simple; it consists of detecting the specific radio waves that were originally generated by a transmitter and then reflected from an obstacle. Radar may be compared with a searchlight that emits a beam of light into the sky. When a flying object is struck by the beam of light, it reflects the light waves back toward the searchlight operator. The operator is then able to observe and follow the object.

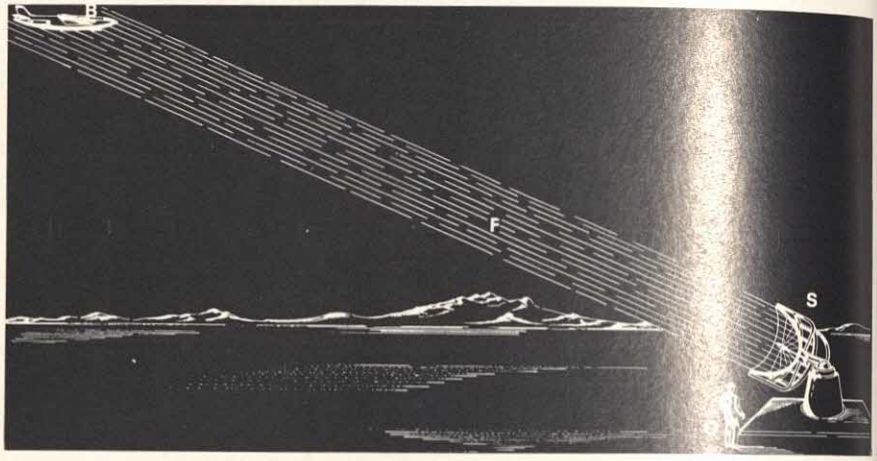
A radar unit emits a beam of microwaves instead of light waves. It has an advantage in that a single piece of equipment contains both the transmitter and the detection system. The radar unit can also plot the movements of an object and its distance from the transmitter.

Radar was employed by many nations during World War II. Radar developments during the five years of the war exceeded all efforts made in the entire field of electronics during the previous 25 years. The techniques perfected in making radar a practical system have subsequently been applied in many other branches of electronics.

Quite apart from its military uses, radar has innumerable applications in meteorology and in marine and air navigation. Radar serves to locate obstacles, detect storm clouds, measure altitude, determine position, and provide a view of terrain beneath cloud banks. Efforts are now under way to develop radar systems that will provide a view through any bad weather formation. The goal of this work is to produce an image similar to the image normally seen by the human eye.

Accurate navigation and control during a space flight to the moon involves radar. On Earth, the tracking stations use two tools, radio and radar, to gather information to be fed into computers that calculate the position and motion of the spacecraft. In California, Australia, and Spain, steerable antennas 85 ft (about 26 m) in diameter that swivel on fixed

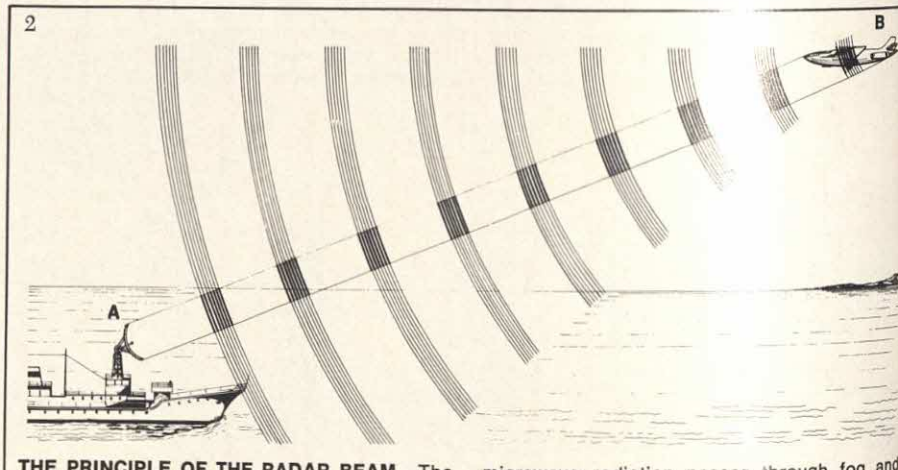
1



THE PRINCIPLE OF THE SEARCHLIGHT— This drawing illustrates the operating principle of a searchlight. **S** is a very intense source of light, enclosed in a system of lenses and mirrors that produce a parallel and concentrated beam of light **F.** If an object **B** enters this light beam, it is struck by the rays of light and reflects them in all directions. A portion of this reflected light returns toward the searchlight. By perceiving this reflected light the observer **O** is able to identify the object.

The difficulties and limitations of the searchlight are immediately obvious. The target **B** is clearly visible only if the background of the sky is dark. Consequently, a searchlight is useful only at night. Furthermore, light rays are extremely sensitive to the presence of mist and the drops of water that form banks of fog or clouds. Under such conditions the searchlight is ineffective.

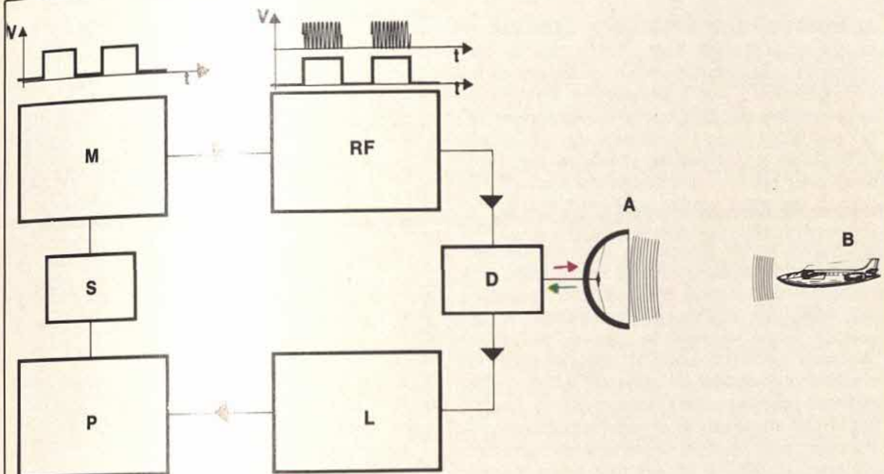
2



THE PRINCIPLE OF THE RADAR BEAM—The system diagramed here is analogous to the searchlight system, but the beam consists of microwaves. **A** is the source—an antenna that radiates energy in the form of microwaves, electromagnetic radiation having wavelengths of a few cm (about 1 in.). These waves are emitted in brief pulses that eventually strike the target **B,** which then reflects them. A portion of the radiation thus returns toward the antenna, where it is intercepted. Microwaves travel at the speed of light. However, the time they take to travel to the target and back to the antenna can be measured, and the distance to the target can be calculated. This system has two advantages over the searchlight: (1) microwave radiation passes through fog and clouds very easily, and (2) the environmental background (sky, ground, sea) does not emit radiation of this type. Consequently, the target always stands out.

Radar makes use of different wavelengths, depending upon the particular application. Some radar installations operate on wavelengths of 10 cm (about 3.9 in.) or more. Most units, however, are designed to operate on wavelengths of about 3 cm (about 1.2 in.). Systems operating on wavelengths of about 1 cm (about 0.4 in.) have also been produced. Finally, certain experimental systems even make use of wavelengths close to 1 mm (about 0.04 in.).

3

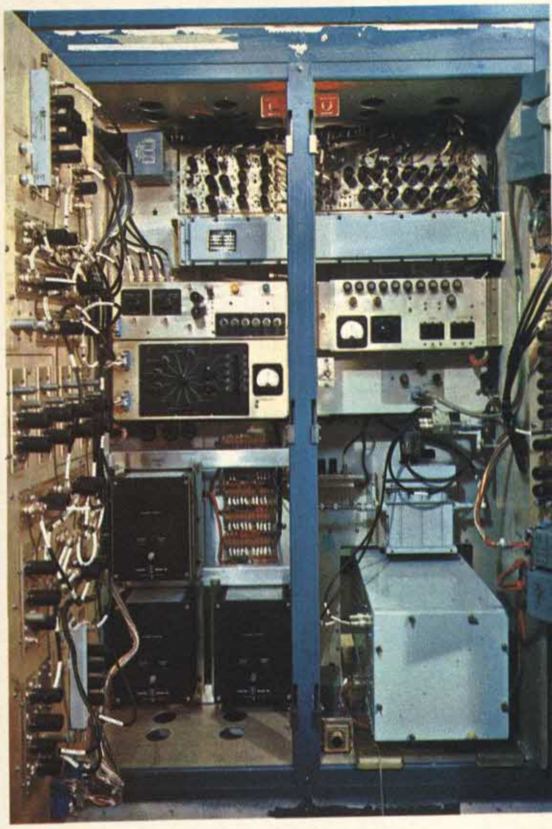


SCHEMATIC DIAGRAM OF A RADAR INSTALLATION—This layout shows the various electronic circuits used in a radar installation. **RF** represents the radio frequency oscillator that generates the high-frequency wave pulses. These are transmitted to the antenna and thence radiated into space. The oscillator operates intermittently in order to emit microwave pulses into space at regular time intervals. This is achieved by means of the circuit **M,** which generates rectangular pulses whose duration (0.01 microseconds in this example) determines the interval of time during which the oscillator sends its pulses to the antenna. The diagrams above the **RF** block indicate that, upon arrival of the pulse sent by **M,** the oscillator emits a series of 10 successive waves (based on a frequency of 1,000 megacycles per second, although much greater frequencies are used in actual practice). These waves follow the path indicated by the arrows and pass through the duplexer **D,** which acts as a shunting device. When the modulator permits the generator to produce alternating radio frequency voltages, the duplexer allows the voltage to pass only in the direction of the red arrow. The voltage is then radiated into space by the antenna.

The emitted microwaves strike the target **B;** a portion of the reflected waves then returns to the antenna, which in the meantime has ceased to emit. In this situation the duplexer reverses its operation and the energy gathered by the antenna is sent in the direction of the green arrow. The receiver **L** collects the signal and converts it into pulses suitable for measurement. Finally, these pulses are displayed by the apparatus **P. S** synchronizes the modulator with the display. It also determines the cadence with which the microwave pulses are emitted by the antenna.

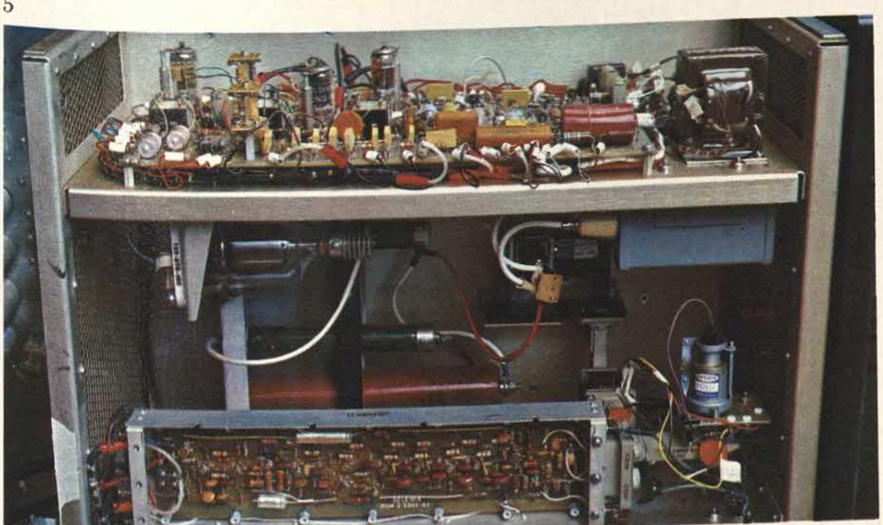
Radar is based on a principle that requires an enormous emission of energy. In fact, using a wavelength of 3 cm (about 1.2 in.) and an antenna having a diameter of 4 m (about 13 ft), a beam that diverges at an angle of one half degree can be produced. At a distance of 100 km (62 mi) from the antenna this beam will be spread out over a circle having an area of about 500,000 m^2 (over 5 million ft^2). A small aircraft, however, has an outline with an area of about 10 m^2 (about 108 ft^2). Thus, a very small part of the emitted energy is reflected by the aircraft, and only about one billionth of the emitted energy returns to the antenna. In other words, both a very powerful transmitter and a highly sensitive receiver are needed if a radar installation is to be practical.

4



A COMPLEX RADAR RECEIVER — Powerful radar installations, such as those used for air traffic control and navigation, are highly complex. This photograph shows the components of the receiving unit of one of these radar installations.

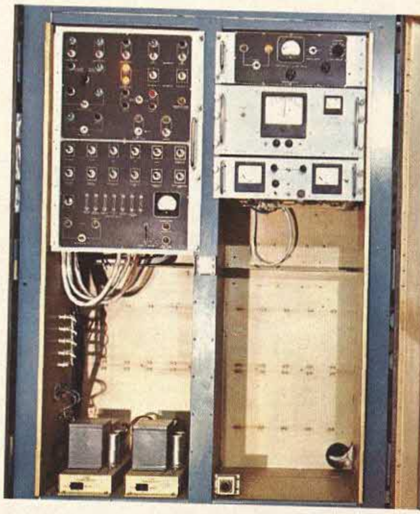
5



THE CIRCUITS OF A SMALL RADAR SET — Certain radar sets are designed to perform supervisory or controlling functions. Because these functions are less important than others, such sets are made up of comparatively few electronic components. The circuits shown here comprise the receiving, transmitting, and modulating units of such a set. In order to display the data, a cathode ray tube circuit is also needed.

The antenna is generally placed at some distance from the electronic circuits of a radar set. These circuits are usually housed, while the antenna is exposed to the weather.

6



AUTOMATIC CHECKOUT AND CONTROL— Large radar installations, such as the one shown in Illustration 4, must not be permitted to break down, for a breakdown may cause considerable damage and even endanger the lives of many people. A special device is used to maintain a continuous check on the entire electronic system of a radar installation. The lighted indicator lamps shown in this photograph indicate proper functioning of these electronic units. Measuring instruments permit more delicate checks to be carried out; instruments are used, for example, to ensure that each part of the equipment receives the correct power supply. In certain cases the control device consists of several identical circuits, only one of which is actually in operation at any given time. Whenever an operating circuit shows signs of reduced efficiency—although still working correctly—the operator replaces it with a circuit that is in perfect working order.

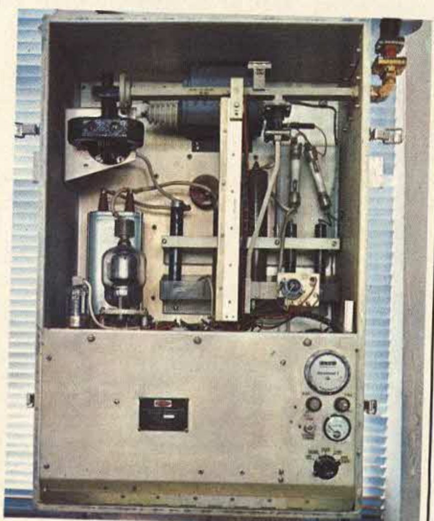
bases help to pinpoint the direction of the radio signals and make radar measurements of the distance and speed of the spacecraft. Information from the tracking stations and computers on Earth is transmitted by the control center to the navigator in space. The ground system is more accurate than on-board navigation out to the distance of the moon.

Radar systems today are rather heavy and costly. The radio frequencies needed for radar can be produced only by means of heavy electronics systems that consume large amounts of energy. However, portable radar sets, produced at low cost and containing semiconductor circuits, have been produced.

This article explains the operating principles of radar.

THE RADIO FREQUENCY GENERATOR— In this receiving, transmitting, and modulating unit, the magnetron tube that produces the radio frequency energy pulses is clearly visible. The connections between the various components are made by means of shielded cables. Each section of cable conveys voltages that must not interfere with others. If the cables were not adequately shielded, the various circuits would not be separated from each other. Interference between circuits could lead to operational malfunctions and could prevent the transmission and reception of signals. A shielded cable consists of a conductor completely covered with insulating material. A cable may also be protected by a conducting sleeve that is connected to ground.

7



8

THE ANTENNA—The microwave pulses are emitted by an antenna located inside the installation. The pulses are then channeled to an external reflector by means of a wave guide, which consists of metallic tubes that are rectangular in cross section and that are polished and gilded on the inside. The wave guide conducts the electromagnetic radiation as far as the diffusion horn (on the right). The waves issue from this horn and strike against the large parabolic reflector (left). The wave pulses leave this parabolic reflector in the form of a flat beam and then follow a straight-line path through space. In actual practice the beam tends to spread out as a result of diffusion. The shorter the wavelength and the greater the width of the antenna, however, the smaller the diffusion.

A radar installation that is intended to pinpoint targets with great precision must have a very large antenna. The antenna can reflect the microwaves even if its surface is not continuous. Here the surface of the parabolic reflector is covered with a dense gridwork for which the size of the individual meshes is related to the wavelength of the radiation being used. The mesh has less wind resistance than a solid sheet of metal.

When the goal is to localize a target with great directional accuracy but not elevation (as is the case when a ship is to be located) the best antenna to use is quite wide but not high—such as the one shown in this photograph. Such antennas must be rotated in order to explore the entire horizon. This creates the problem of manufacturing a rotating joint capable of turning despite weather conditions, as well as special articulation in the wave guide.

RADAR—II | applications of the microwave eye

Radar is so widely used in so many different ways today, from astronomy to rocket guidance and detection, that few remember how young it is. Furthermore, perhaps fewer still realize that its development was greatly stimulated by World War II and that it played a vital part in the outcome of that war. If England's Royal Air Force had not had radar to alert fighter squadrons as to where the Luftwaffe would strike next and in what strength, the Battle of Britain might have been lost and America's later contributions might have been of no avail.

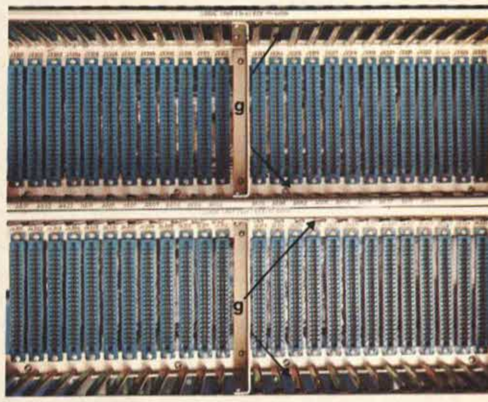
Radar had its origins in the discovery of radio waves and the later developments of Marconi, de Forest, and others in the early years of the present century. Its operational principle is simple—that a radio wave, like a sound wave, is reflected by any object it strikes. Radar equipment is capable of a great variety of applications, and its construction varies in complexity according to the application. For example, a radar installation may simply be required to detect the presence of an object ahead, while at other times it may be necessary to have an image of the area or object being detected. The equipment needed for the latter purpose is far more complex than that needed for the first.

The simplest radar is one that emits radio waves in a fixed direction. If the emitted signals encounter an object, they are partially reflected and perceived by the radar's receiver unit. Such equipment merely indicates the presence of an object, but even it requires electric circuits of a very complex and multipurpose nature.

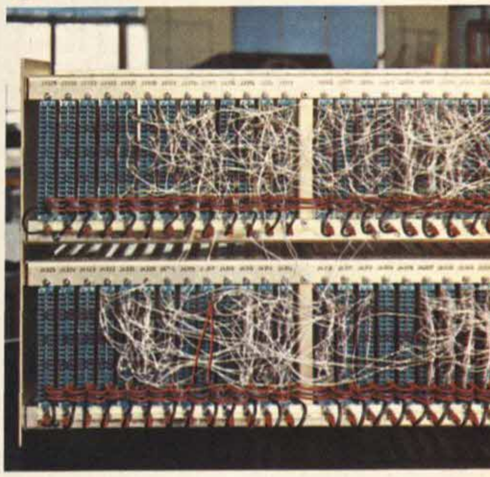
Such equipment first of all enables accurate determination of the direction in which the object is detected. The direction in which the antenna points

1a

1b



1c



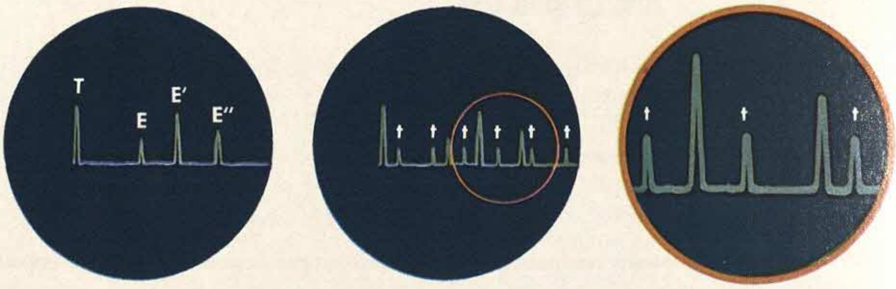
THE COMPLEX CIRCUITS OF A RADAR SET —Some circuits in a radar set are designed to give high power outputs, while others are designed to give small outputs. The former circuits are the ones that generate the pulse to be transmitted. This pulse may have a power of many kilowatts even in a small radar set, but in long-range radar it may reach a power of thousands of kilowatts. The other circuits are made up of components of lesser power and serve to refine outgoing pulses, obtain samples of these pulses, stabilize the frequency, filter the echoes, control the direction of the presentation, and so forth. The number of these components depends on the complexity of the functions they have to carry out. The set must, therefore, be assembled in such a way that the various circuits can be readily removed in order to be checked and, if necessary, quickly replaced. This is achieved by the extensive use of printed circuits.

Photographic techniques are used to print the circuit on a base made of insulating material. The base is usually called a board, and the copper lines printed on it connect the various components of the circuit. The components are soldered onto the board according to the circuit requirements. Illustration 1a shows a printed circuit completely assembled.

The circuit board is then inserted in a container made specifically for the purpose, such as the one shown in Illustration 1b. The terminals of the board are slipped into the metal clips at the bottom of the container, while the board itself is held in place by the lateral guides indicated by the letter **g**. Trouble caused by a defective board, which can be identified by the use of a checking device, can thus easily and quickly be rectified by replacing the board in question. The wires connecting the various boards are at the back of the container. Illustration 1c shows the maze of wires (white) that connect the terminals to one another and wires of other colors (red, purple) that carry power to the boards or ground some element of the circuit.

2

a b c



THE TYPE A PRESENTATION—The simplest radar model determines the position and distance of a target. The measurement of distance can be achieved rather simply by showing both the emitted pulses and the received pulses on the screen of an oscilloscope. This type of visualization of return data, which is technically the simplest way of signaling objects, is called Type A presentation. The first radar echoes from the moon, received in 1946, were obtained by means of a Type A presentation.

The luminous spot of an oscilloscope is made to travel across the screen in the time required for the radar beam to travel the maximum distance the instrument is capable of sending it and to return in the event it is reflected by an object. For example, if a radar has a range of 150 km (about 90 mi), the reflected radiation will return in 0.001 sec. The luminous spot will, therefore, sweep the screen of the oscilloscope in 0.001 sec, starting its sweep every time the radar emits a pulse. At the same time the pulse is emitted into space, another pulse is sent to the oscilloscope so that the trace is deflected upward and forms the peak **T,** which marks the starting point for the measurement of times and distances. From this moment on, whenever the antenna receives a return signal, another pulse will deflect the trace and form the peaks **E, E′,** and **E″,** as shown in Illustration 2a. The interval separating these echo peaks from **T** is a measure of the distance of the objects from the antenna, while the height of the peaks indicates the dimensions of the corresponding objects.

The presentation system shown on this screen in Illustration 2b is more complex than the first, for here the screen of the oscilloscope registers both the signals received by the antenna and periodic signals emitted by a special oscillator. These oscillator pulses have a special form and a constant height; they represent a time scale (**t, t, t,** and so forth) on the oscilloscope screen against which the echo return times (and therefore the distances of the objects) can be measured.

Certain radars are used for very accurate measurement of distances. They function in much the same manner as the radar shown in Illustration 2b, but the time scale on the screen is enlarged so that the screen shows only that part of the scale corresponding to the distance at which the object is expected to be found (Illustration 2c). The screen shown in Illustration 2c also registers the peaks due to the return signals received by the antenna.

3

TYPES OF SCANNING—The greater the field a given radar is capable of scanning, the greater its usefulness. The form of that field depends on the uses for which the radar is intended.

The type of antenna shown in Illustration 3a is used for scanning the horizon at sea. The antenna is very wide but has a limited height. It is highly directional horizontally, because it is important to distinguish between two targets that are close together, but it is much less directional vertically because vertical discrimination is less important.

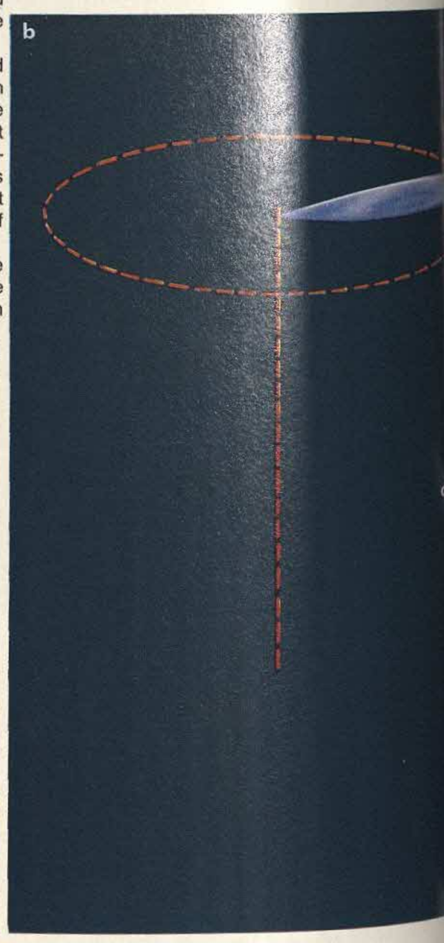
One type of scanning is the simple sweep of the horizon (Illustration 3b). It is used in most radar applications and particularly in radars used for marine navigation, like the one shown in Illustration 3a, where the field to be explored does not extend above the horizon.

A second type of scanning is combined horizontal and vertical scanning (Illustration 3c). This is accomplished by rotating the antenna slowly about its vertical axis and at the same time moving it rapidly a short distance up and down. This type of scanning is used for the detection of aircraft, because it is important to know not only the position of the aircraft but also its altitude.

A third type of scanning is used to explore the space within a limited cone. In this case the radar beam scans along the spiral shown in Illustration 3d. Radar of this type is often

a

b



indicates the direction of the object, provided the antenna both emits and receives radio waves.

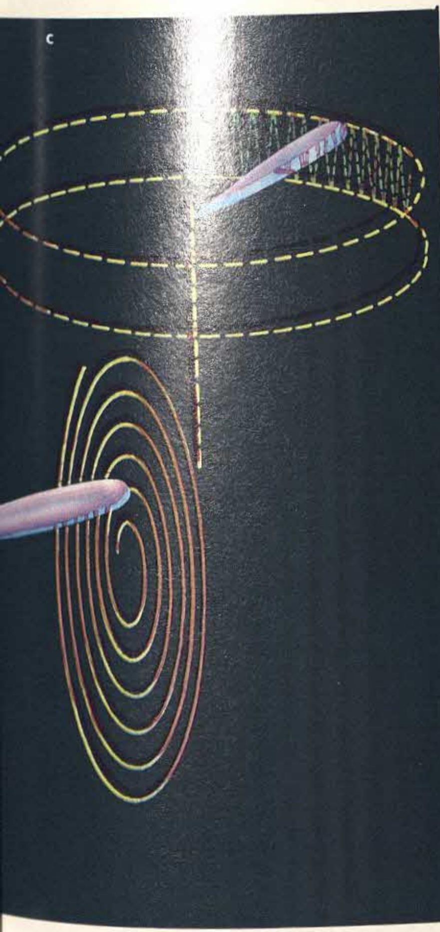
Furthermore, by measuring the elapsed time between the transmission of the signals and their return, it is possible to determine the distance of the object. Radio waves travel at the speed of light and therefore cover about 300,000,000 m/sec (about 186,000 mi/sec). Thus the waves reflected from an object 15 km from the installation will return to it about 0.0001 sec later. Since electric circuits can measure intervals of time as brief as one billionth of a second, it is possible to equip a radar installation so that it can measure distances accurately. It is also possible to measure the speed with which an object is moving. The waves reflected from an object approaching or moving away from the antenna will not have the same frequency as the waves that were emitted. This phenomenon is known as the Doppler effect. If the object is approaching, the frequency will be higher; if it is moving away, the frequency will be lower.

4

THE ARTICULATED JOINT OF THE WAVE GUIDE—A component that plays an important role in radar scanning is the articulated joint of the wave guide that channels the microwaves to the antenna. Shown here is a rotating antenna; below it is the rising part of the wave guide **v**, which then continues horizontally **o**. The two parts are connected by a delicate rotating joint that ensures that the waves inside the wave guide will be emitted regularly while the antenna is rotating.

used aboard aircraft to detect objects ahead.

c



5

THE TYPE PPI PRESENTATION—A practical system of showing the echoes received by a radar is the so-called PPI system (from the initials of the words "plan position indication"). This system operates on the principle shown in Illustration 5a.

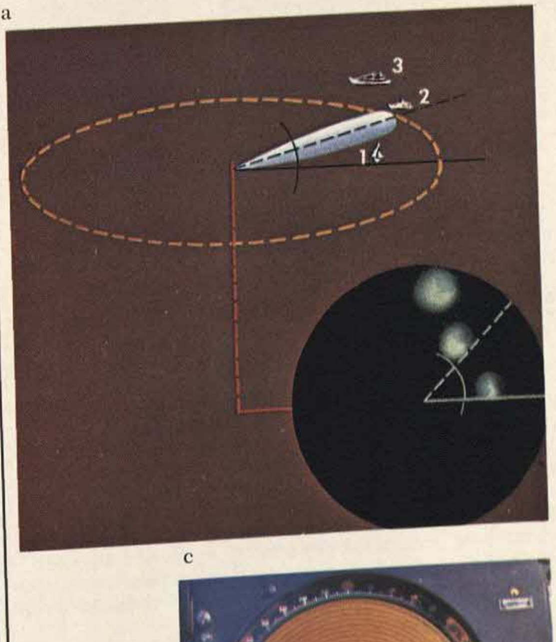
The radar beam explores the space around the antenna and encounters various objects that reflect the echoes 1, 2, and 3. Not all the echoes return at the same time. Some arrive after a certain interval, others later according to their distances from the antenna. At the same time an electronic beam is made to sweep along a radius of the screen of an oscilloscope; this beam has a very low luminosity but becomes bright as soon as an echo is received by the antenna, the brightness being in direct proportion to the intensity of the echo itself. Bright spots will, therefore, form on the screen, their positions along the radius corresponding to the distance of the objects that cause the echoes. The beam sweeping the screen is synchronized to rotate with the antenna; thus the whole field of the screen gradually shows traces or images of the echoes. One important characteristic of the screen is the fact that the images created on it by the echoes will last until the beam returns to sweep the same point again.

Illustration 5b shows the screen on which the radar echoes appear. The radar in this particular case has a medium range. Other radars have a range of many thousands of miles and have screens much larger than the one shown here.

Illustration 5c shows the screen of a meteorological radar in operation. Such a radar is specially adapted to receive echoes caused by clouds, hail, and snow. It is used in airborne navigation as well as in meteorology.

Illustration 5d shows the unit that supplies the high-voltage current needed by the oscilloscope tube of the radar. It consists of a transformer that raises the network voltage from 115 V (volts) to 13,000 V and of a rectifier that converts alternating current into direct current. The circuits on the right stabilize the voltage output of the unit.

a



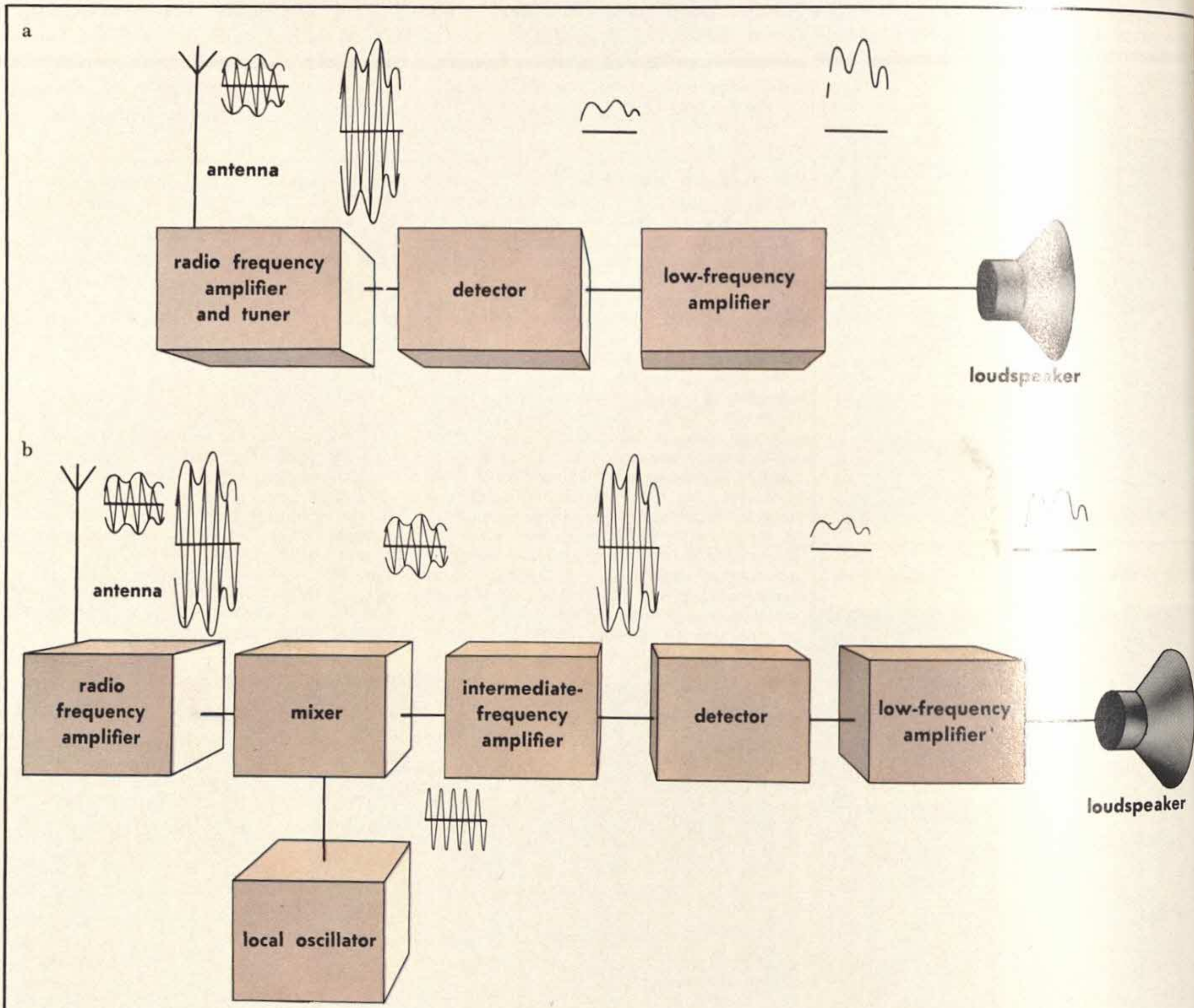
b

c

d

RADIO—I | frequency modulated superheterodyne diode receivers

1



RADIO RECEIVER LAYOUTS—Direct amplification receivers, like that shown in Illustration 1a, consist of the following elements:

1. An antenna for intercepting radio waves;
2. A radio frequency amplifier and tuner, directly connected with the antenna circuit, which by means of variable capacitors in the oscillating circuits becomes resonant with the desired reception frequency and then amplifies this frequency in one or more stages;
3. A detector that separates the original audio signal from the high-frequency carrier;
4. A loudspeaker that converts the audio signal into sound waves by a process that is the reverse of the one used in a microphone.

All present-day commercial radio receivers are based on the superheterodyne circuit shown in Illustration 1b and consist of the following elements:

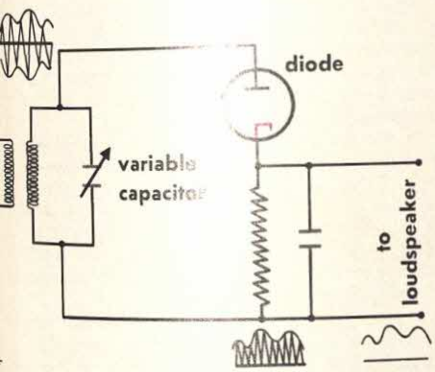
1. A radio frequency amplifier that amplifies the entire range of received waves but does not tune;
2. A local oscillator that by means of a variable capacitor generates a variable frequency having a fixed difference in relation to the desired reception frequency;
3. A mixer that unites the incoming frequency signal with the locally generated signal and supplies an intermediate output frequency, equal in value to the difference between the two, which is modulated by the incoming signal;
4. An intermediate frequency amplifier that amplifies the intermediate frequency signal;
5. A detector that separates the original audio signal from the intermediate frequency carrier;
6. A sound frequency amplifier and a loudspeaker.

In a manner of speaking, a radio receiver is a radio transmitter in reverse. That is, it is a combination of electronic circuits capable of intercepting the electromagnetic waves radiated by a transmitter and converting them into audible sounds. However, its combination of circuits begins with the element with which the transmitter ends, the antenna.

Theoretically at least, the antenna of a radio receiver is identical with the transmitting antenna. It is somewhat simpler in actual practice, but it always consists of an oscillating circuit that resonates at the frequency of the waves to be intercepted. These waves generate induced electromagnetic forces and consequently oscillating currents in the an-

THE DIODE RECEIVER — This is the simplest type of radio receiver. It is derived directly from the crystal receiver used in the earliest days of broadcasting and is based on the same principle. The antenna is connected to the primary winding of a transformer whose secondary winding forms part of an oscillator circuit. This circuit contains a variable capacitor that enables it to be tuned to the desired frequency. The radio signals intercepted by the antenna induce electromotive forces in this oscillator circuit. These forces or potentials attain substantial value only in the case of the resonance frequency. The oscillator output is connected to a diode (a galena crystal would do the same job) and this allows only the positive voltage to pass; that is, one half of the received radio frequency signal.

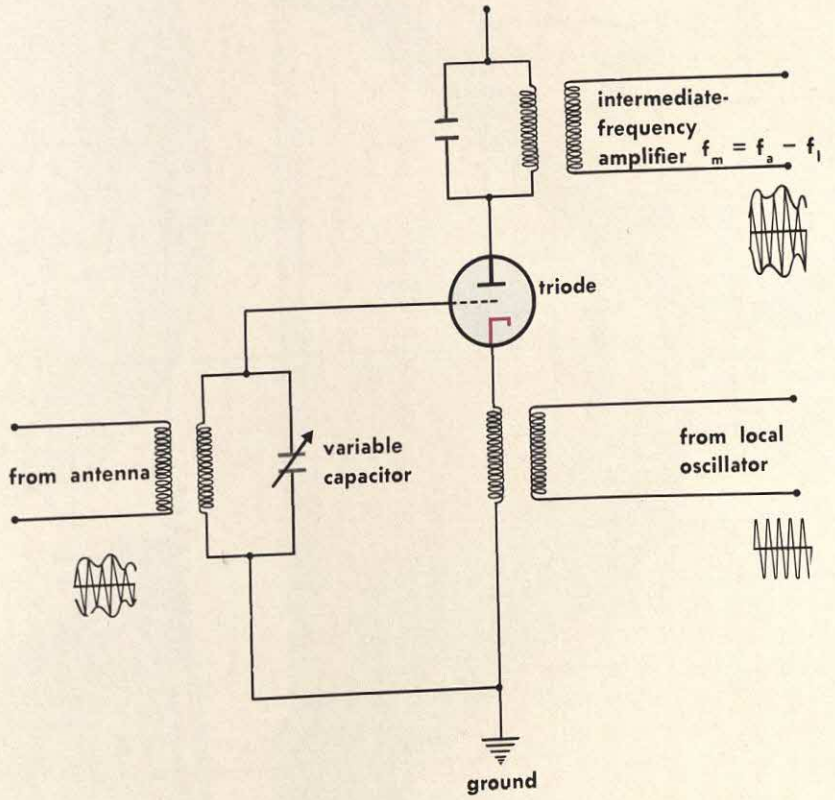
2



These oscillations are then applied to the terminals of a circuit consisting of a capacitor and a resistor placed in parallel. The voltage passing through the diode charges the capacitor; then, while no current is passing through the diode, the capacitor discharges across the resistor. During the next positive half-wave the diode once again conducts current, the capacitor becomes charged, and the cycle is repeated. As a result, the capacitor terminals, to which the loudspeaker is connected, carry a continuous voltage corresponding to the peak value of the radio frequency voltage.

However, this voltage is modulated; that is, it varies slowly in relation to the carrier frequency. Therefore, the capacitor voltage changes as the peak values change, and the capacitor' output voltage thus corresponds to the audio signal transmitted to the receiver.

3



THE SUPERHETERODYNE RECEIVER — The characteristic feature of the superheterodyne receiver is that it produces a local oscillation that, when subtracted from the carrier frequency to be received, has a fixed difference known as the intermediate frequency. For example, if the carrier has a frequency of 1,500 KHz, the local oscillator must produce a frequency of 1,675 KHz in order to obtain a fixed difference or intermediate frequency of 175 KHz.

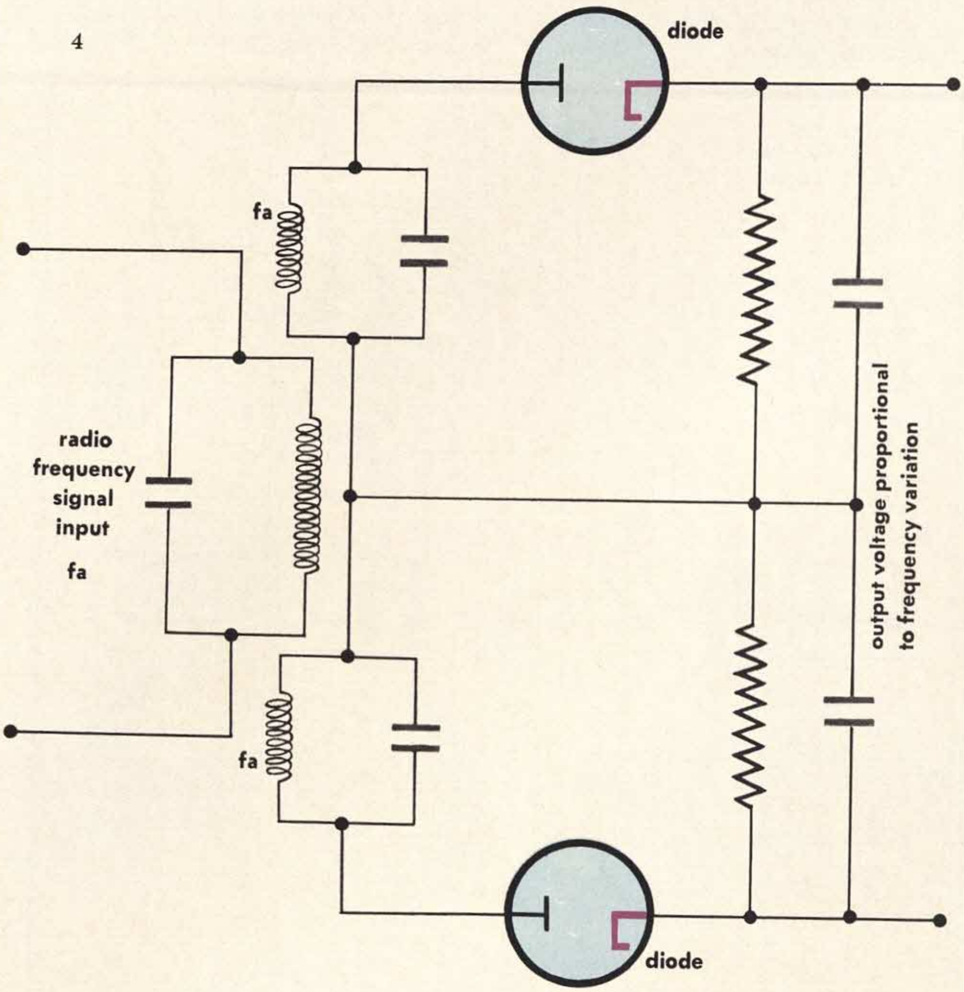
The superheterodyne circuit has a considerable advantage over other circuits because, with the exception of a first stage in high frequency, it produces all amplifications at a fixed frequency value and therefore makes possible better selectivity and fidelity.

Illustration 3 shows in schematic form a simple type of mixer stage, the most important part of a superheterodyne receiver.

The signal received by the antenna is carried to an oscillator circuit on the input side of the mixer stage. The variable capacitor tuning this circuit is connected to the condenser of the local oscillator in such a way that both capacitors simultaneously vary the resonance values of their respective circuits, and a fixed difference of 175 KHz is thus preserved over the entire range of frequencies. The signal is applied to the grid of a triode, or vacuum tube, and is thus added to the voltage of the frequency produced by the local oscillator, which is applied to the cathode of the tube.

The plate circuit of the triode thus carries a complex signal that is the sum and the difference of the applied input frequencies. Usually only the difference is used. A suitable resonance circuit, permanently tuned to this difference of 175 KHz, carries the intermediate frequency output signal, which can subsequently be amplified and demodulated.

4



FREQUENCY MODULATION RECEIVERS—Radio broadcasts are sometimes made by means of frequency modulation signals. Frequency modulation has many advantages, particularly in fidelity and lack of static interference. A receiver designed to receive frequency modulated or FM signals differs from an ordinary receiver for amplitude modulated or AM signals in that the FM signal must first be converted into an AM signal. This signal is then amplified just as in a superheterodyne receiver.

The conversion is accomplished by means of a so-called discriminator circuit, which transforms frequency variations into variations of the output voltage of the circuit.

One of the simplest methods is that of the balanced discriminator, which consists of two resonance circuits magnetically coupled with the resonance circuit at the end of the antenna stage. These circuits are set to two different frequencies, one slightly less and the other slightly greater than the carrier frequency f_a of the signal to be detected. Each circuit contains a diode with the cathode connected to the output terminal. When the discriminator circuit receives the unmodulated carrier signal, both its resonance circuits oscillate at a frequency equidistant from their respective resonance frequencies and give two equal signals. The diodes, therefore, balance each other and no output signal for the complete circuit is produced.

If the frequency varies as a result of the modulation, the frequency will come closer to one of the resonance frequencies, the circuit in question will oscillate more strongly, and its diode will supply a greater voltage than that of the other circuit. The opposite result will be obtained if the input frequency varies in the other direction.

The output voltage of the discriminator circuit is thus proportional to the frequency variations of the waves received by the antenna.

tenna. Some antennas are highly selective; that is, they are especially sensitive to a particular frequency. This is true of a radio-telephone link, which operates at a constant frequency. Other antennas, however, such as those of radio and television receivers, must be able to receive a wide range of frequencies.

THE TUNING CIRCUIT

Selection of the desired frequency from among the many frequencies an antenna can receive is accomplished by a tuning circuit. Two types of tuning circuits are used, direct amplifying receivers and superheterodyne receivers.

In direct amplifying receivers, which are used only for special purposes, the antenna circuit itself creates different conditions of resonance, generally by means of a variable capacitor in the oscillating circuit. In this way only the frequency that agrees with the conditions of resonance creates an electromagnetic force strong enough to be further amplified.

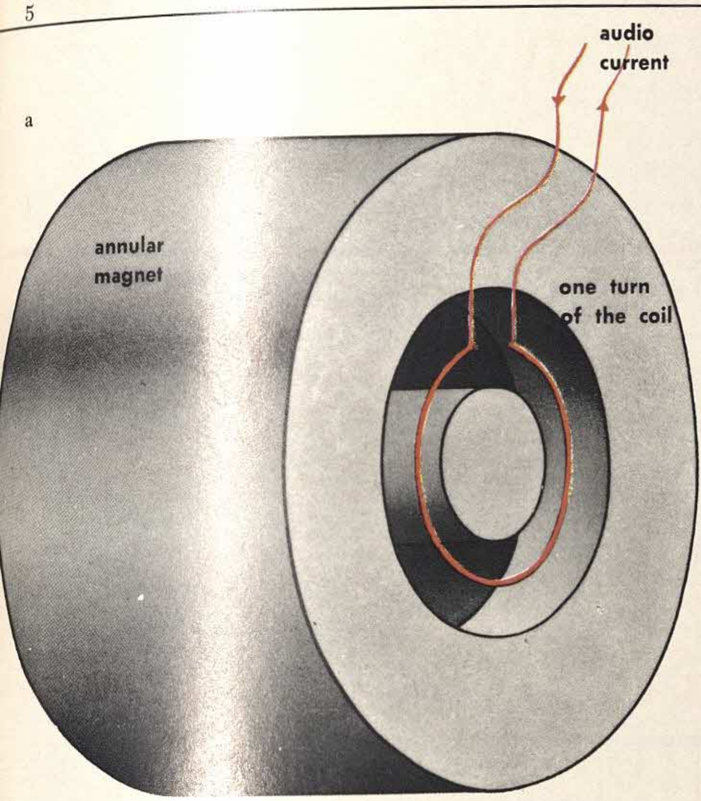
In superheterodyne receivers, on the other hand, the frequency received is mixed with a locally generated frequency, and an intermediate frequency of fixed value is obtained that is transmitted to the later stages of the receiver. Thus in superheterodyne receivers the whole range of radio waves is subjected to high-frequency amplification, and the task of tuning in a single station is left to the mixer stage, where the variable local frequency is introduced.

The next stage for both types of receivers is the detector or demodulator, which eliminates the high-frequency carrier waves and restores the originally transmitted signal.

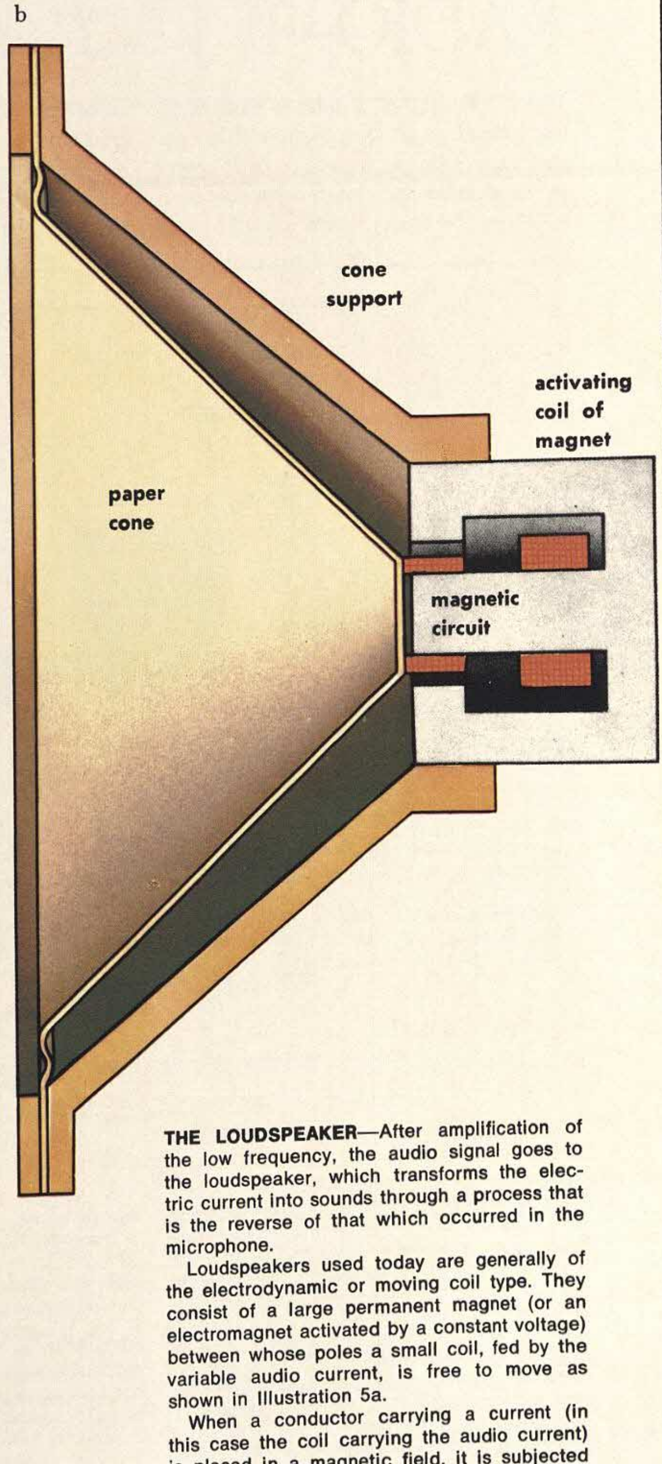
From this stage the signal goes to the power amplifier, which brings it up to the desired reception level and sends it to the loudspeaker. As the sounds come from the loudspeaker, they are heard just as they were picked up by the microphone.

5

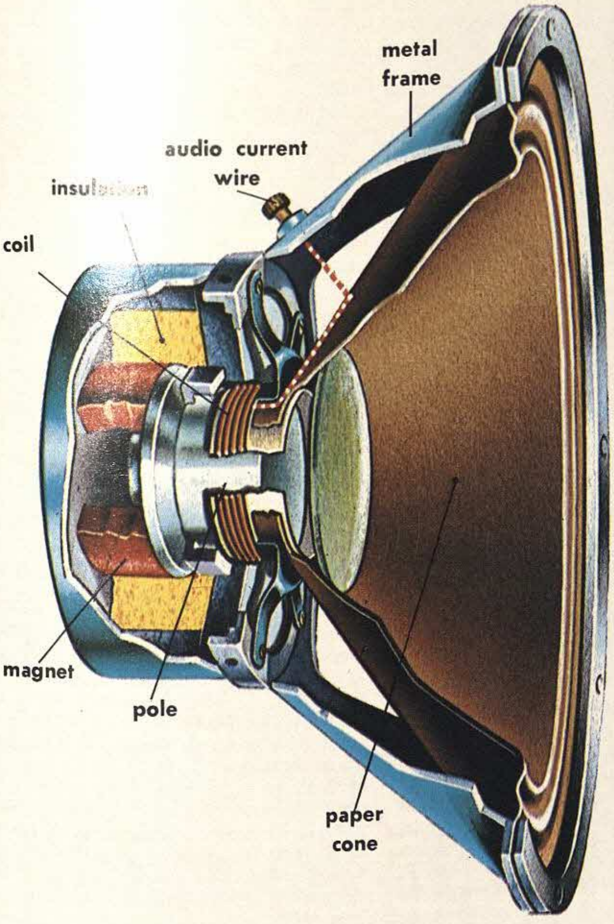
a



b



c



THE LOUDSPEAKER—After amplification of the low frequency, the audio signal goes to the loudspeaker, which transforms the electric current into sounds through a process that is the reverse of that which occurred in the microphone.

Loudspeakers used today are generally of the electrodynamic or moving coil type. They consist of a large permanent magnet (or an electromagnet activated by a constant voltage) between whose poles a small coil, fed by the variable audio current, is free to move as shown in Illustration 5a.

When a conductor carrying a current (in this case the coil carrying the audio current) is placed in a magnetic field, it is subjected to a mechanical force perpendicular to both the electric and the magnetic fields. Since the current is continually changing in strength, the coil is forced to oscillate along its axis.

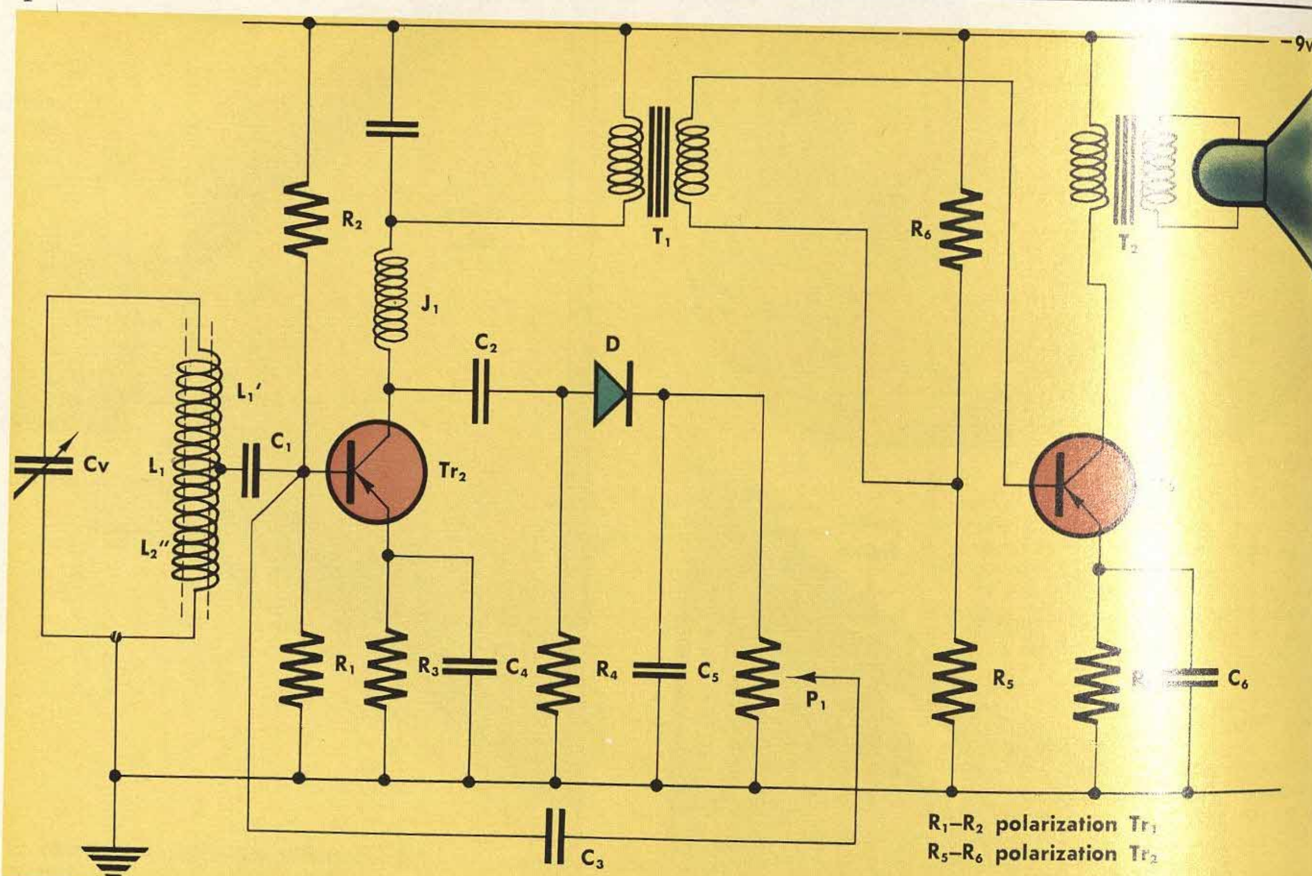
Illustration 5b shows a schematic section of a loudspeaker, while Illustration 5c shows a sectional view of a loudspeaker. Both views show the paper cone.

RADIO—II | a simple transistor radio

The functioning of a radio is most easily understood by an examination of its construction. A simple transistor radio serves as the example here. Although simple, it contains the major parts of any radio receiver: the radio-frequency (high-frequency) part, the detector, and the audio-frequency (low-frequency) part.

The radio-frequency part receives radio signals from the broadcasting station, amplifies them, and passes them through the detector to the audio-frequency part. The detector converts the radio signal to a single (direct current) audio signal. The audio-frequency part amplifies this

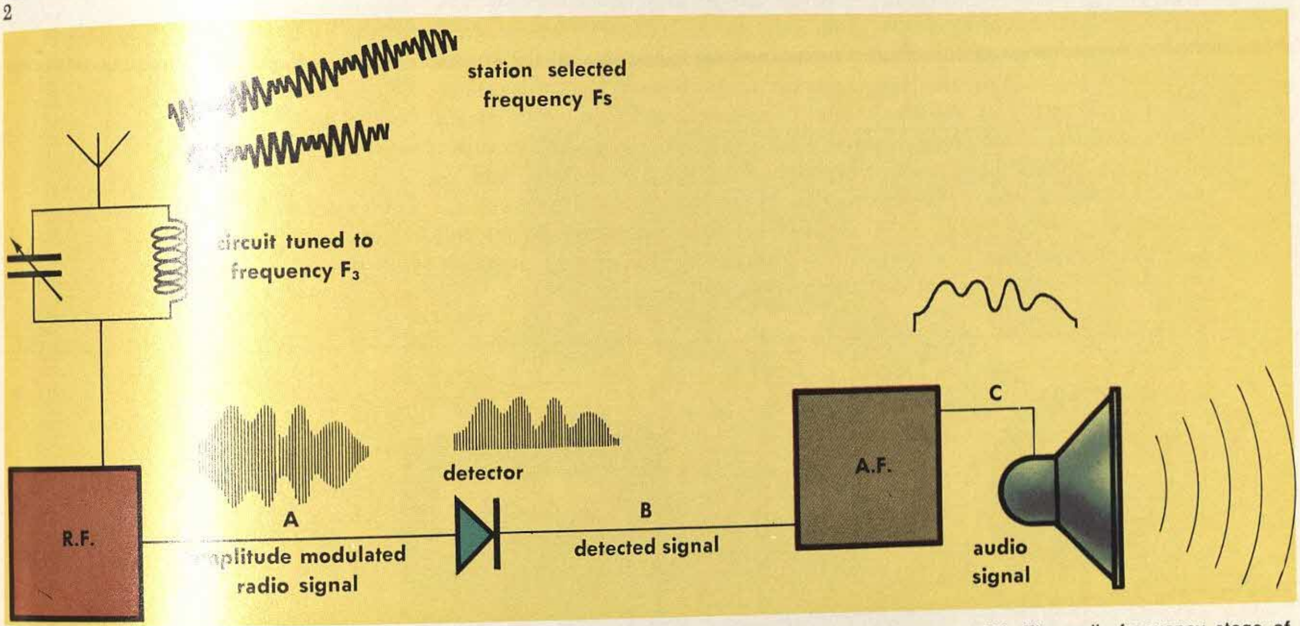
1



REFLEX RADIO RECEIVER — This illustration is a diagram of the circuit of a simple transistor radio. The radio signal enters from the left side of the circuit displayed, which represents the tuned circuit formed by a ferrite antenna wound with litz wire. Here the radio frequencies induced by the antenna are initiated. Each frequency can be precisely selected from within a certain range of frequencies and monitored. The frequency selected is determined by components **L_1** and **Cv**, a coil and variable capacitor, respectively. These two components, **L_1** and **Cv**, mounted in parallel, form a circuit tuned to a frequency that may be varied by adjusting the position of **Cv**. If **Cv** is adjusted to a frequency equal to that of a broadcasting station, the signals from that station are received. Coil **L_1** consists of two sections: **L_1'** and **L_2''**. The terminals of **L_2''** produce the radio signal that has to be amplified, detected, and converted into an audio signal.

The signal from the coil is conducted through capacitor **C_1** to transistor **Tr_1**, which represents the radio-frequency amplification stage. This transistor also has another function. The signal enters **Tr_1**, which is biased to a given value by resistances **R_1** and **R_2**. It is amplified as a radio-frequency signal and arrives at the output of this transistor's collector. Because it is a high-frequency signal, it can now only pass through capacitor **C_2**, not through inductance **T_1**. This is because high frequencies meet strong resistance in passing through inductances, while capacitors put up little resistance. The opposite is true for low frequencies, which do not pass through capacitors, but are almost unaffected by inductances. Thus, the high-frequency signal leaving the collector can pass only through **C_2** and enter diode **D**. This is the detector diode that changes the frequency from a high to a low frequency (from a radio to an audio frequency). The signal then arrives at potentiometer **P_1**, where one part of it is isolated and used to regulate the volume. The signal leaving **P_1** arrives at **Tr_1** by passing through capacitor **C_3**, which must have a high capacitance in order to let low frequencies through. When the signal reaches **Tr_1**, it is amplified again. It leaves the collector as a low-frequency signal that passes through **J_1**, but not **C_2**. The low-frequency signal leaving **Tr_1** goes to the primary of transformer **T_1** (coupled transformer) and is transmitted from the secondary of **T_1** to transistor **Tr_2**, where it is amplified again. It is then dispatched to output transformer **T_2**, which controls the loudspeaker. The radio-frequency stage consists of **L_1, Cv,** and **Tr_1**, while all the other components belong to the audio-frequency stage. Transistor **Tr_1** serves both stages, however.

2



ANTENNA AND DETECTOR STAGE—Before radio waves can be detected, amplified, or modified, they must be picked up by the receiver. This is done by the antenna. The antenna projects into the space wherein the radio waves are traveling and picks them up. The length of an antenna may vary from a few centimeters to many meters, but each radio must have one. In the example, the antenna is a ferrite cylinder, which also acts as a support for the wire that forms the coil of the tuned circuit. Such an antenna is magnetic and is generally suitable for small receivers designed for listening to nearby stations. Because of its simplicity, this antenna performs best when oriented with respect to the station transmitter.

The detector stage is schematized in the illustration. The radio signal induced in the coil has a frequency equal to that of the transmitting station and has a varying amplitude. The signal consists of a fixed frequency characteristic of the transmitting station, on which an audio signal that modulates its amplitude is superimposed. This is called amplitude modulation, or AM. The radio-frequency stage of a receiver amplifies the signal **A.** This signal is sent to the detector diode, which, as a unidirectional conductor, passes on only the positive component in the radio-frequency signal. It rectifies the signal, just as alternating current is rectified to direct current. The signal is then as depicted in **B.** The audio-frequency component then changes the signal to audio signal **C,** which is strong enough to activate the loudspeaker.

signal so that it can activate the loudspeaker.

The detector is the link between the radio-frequency and audio-frequency parts. In the example (the simple transistor radio), the detector is represented by a simple diode. More complex radios have a second detector, and the radio-frequency signal is converted to an intermediate frequency and then to an audio frequency.

The radio-frequency part not only amplifies the signal, but also selects the desired station from all stations available. This capability (called selectivity) is important if the listener is to pick up only one station at a time without background noise.

Another function of the radio-frequency part is sensitivity—the capability of picking up distant stations. The set portrayed in the illustrations has a low sensitivity; it can receive broadcasts from nearby stations only.

The audio-frequency part of the simplified set used as an example merely amplifies the audio signal from the diode. Amplification is necessary because the signal received from the diode is too weak to activate the loudspeaker.

Audio-frequency circuits work by applying an alternating voltage to a vacuum tube or to a transistor. One terminal of the source is connected to the central slice of the transistor. The opposite terminal is connected to the negative slice of the transistor. Amplification occurs because application of a small voltage change produces a large current change in the transistor.

The amplified current is direct, and current from one amplification stage can operate a small loudspeaker. Usually, several amplification stages are used. At each stage the direct current is transformed back to alternating in order to feed alternating voltage to the next slice of the transistor. One stage in such a series is used for detection. This produces an audio-frequency current that is nearly identical with the envelope of the amplified radio-frequency carrier. This is done by passing the current from the transistor through a capacitor and a resistance that are connected in parallel. The condenser can charge as rapidly as voltage is fed to it by the radio-frequency input. It can only discharge, however, through the resistance, and since this is high, the discharge rate is low.

After one or more stages of amplification, the signal goes to a power transistor, which drives the loudspeaker.

RADIO—III | from the broadcasting studios to the home

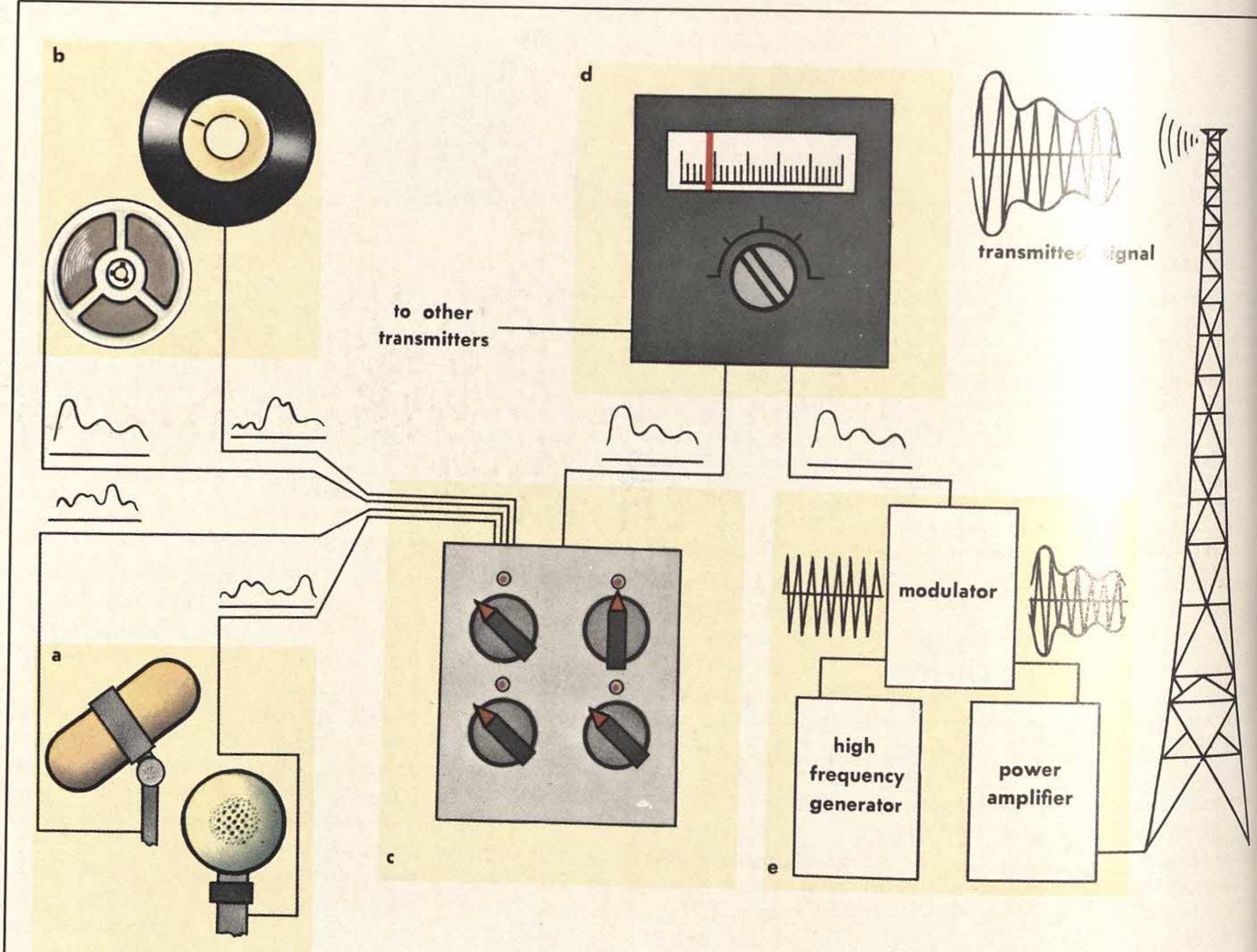
Radio and television, the two most recent methods men have devised for communicating with one another, are both based on one of the most interesting phenomena of the universe—electromagnetic radiation. Electromagnetic radiation comprises a vast range of rays that permeate space as far out as man has been able to probe it. This radiation ranges from cosmic rays of very high frequencies through visible light to infrared rays, which extend to the lowest frequencies.

While infrared radiations are produced by molecular, atomic, or nuclear phenomena, man also has learned to create radiation by means of special electronic circuits. These man-made radiations have frequencies that are much lower than those of infrared radiation, but their range is nevertheless of great importance, for it contains all the known systems currently used for the transmission of messages and pictures by means of electromagnetic waves, such as radio, wireless telegraphy, television, and so forth. These electromagnetic waves are commonly called radio waves. They will move through a vacuum; they do not require a material medium.

HIGH-FREQUENCY CURRENTS

Before examining the different parts of a radio or television transmitting station, one must understand how the conditions needed for transmitting radio waves can be created. Two elements are essential. First, a series of circuits is needed to generate and modulate a high-frequency current. This current is then transferred to other conductors that change the current into radio waves. These latter conductors constitute the antenna; this discussion concerns only the signal at the transmitter.

1



LAYOUT OF A BROADCASTING STATION— Direct broadcasts are made from a sound-insulated studio (Illustration 1a) where the microphones are located.

An adjacent room (Illustration 1b) serves as a recording center. In this room are the tape recorders and record players used for reproducing programs that have already been recorded. All lines of communication are channeled into the control room (Illustration 1c), from which the broadcast is directed. Here the various volume levels are controlled and the signals are mixed.

The output from the control room is the signal to be transmitted. This is channeled to the distribution bars (Illustration 1d) and then distributed to the various transmitters of the network (Illustration 1e) through a switchboard and selectors of the type used in telephony.

At the transmitter the signal modulates the high-frequency carrier current, which is then sent to the antenna that transmits the radio waves.

2a

2b

2c

THE BROADCASTING STUDIO—Broadcasts can be made either by live transmission from microphones placed in the studios or by transmission of programs previously recorded on phonograph records or magnetic tape.

In the case of live transmission, the broadcasting studio is equipped with microphones having characteristics suitable to the particular type of transmission being made. For example, a string quartet (Illustration 2a) requires a microphone capable of gathering omnidirectional sounds, while for a soloist (Illustration 2b), a highly directional microphone is needed to take in only the sounds emitted by the single instrument. Each studio has its own control room, and all the microphones are connected to this room. A control panel enables the operator to regulate the volume levels and to mix the sounds; that is, to channel the sounds gathered by the microphones into the outgoing line. A soundproof glass panel separates the control room (Illustration 2c) from the studio, so that it is possible to see everything that happens in the studio and to issue all orders relating to the broadcast. The first stage of the long road taken by the sounds, which are now transformed into electric signals, takes them to the low-frequency amplifiers, where the modulated currents are strengthened so that the signal may reach its destination without becoming too weak for good reception.

A studio broadcast is generally sent out over a network that makes use of a large number of transmitters. This is the task of the program distribution section, which receives the programs and, using automatic equipment similar to that used in telephone exchanges, distributes them to the various transmitters according to broadcasting schedules.

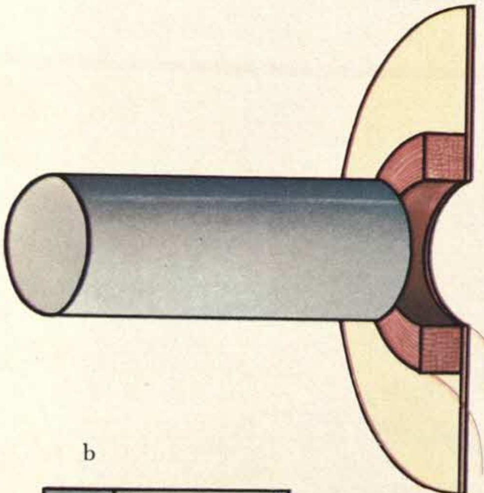
3

MICROPHONES—The first step in sound transmission is the transformation of sounds into variable electric currents, a task performed by microphones. Each microphone consists essentially of a transducer, which is sensitive to sound waves and which vibrates when it is struck by them (Illustration 3a). The microphone is electrically or magnetically connected to the output circuit, in which it causes current variations in direct proportion to the strength of its own vibrations. One type of microphone, the ribbon microphone (Illustration 3b), consists of a thin metal ribbon placed between the poles of a permanent magnet and connected to the electric circuit. The sound waves act on the ribbon, whose vibrations give rise to an electromotive force and consequently to an induced current.

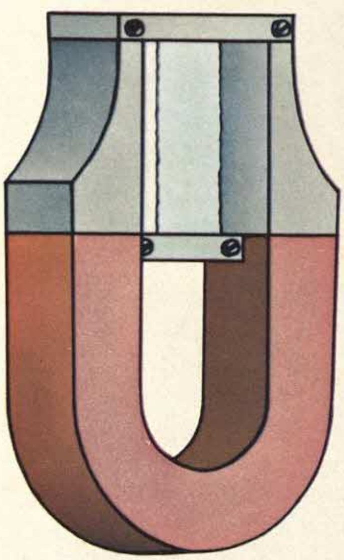
The piezoelectric microphone (Illustrations 3c and 3c′) uses a crystal, made of Rochelle salt or monammonic phosphate, that acts as a transducer. One end of the crystal is fixed, while the other end is connected to a vibrating element. The sound waves thus cause the crystal to vibrate, and as it does it generates an electromotive force that gives rise to a variable electric current. This variable current is always weak and is augmented by a preamplifier placed close to the microphone. Thus, the output of a microphone consists of a current whose intensity and frequency reproduce the volume and pitch of the sound that generated it.

The condenser microphone (Illustrations 3d and 3d′), on the other hand, contains two electrodes, one of which is a fixed plate and the other a membrane that is vibrated by incoming sound waves. The two electrodes are set a few microns apart, and the air between them acts as a dielectric. The vibrations of the membrane cause the dielectric to vary in thickness; this in turn varies the capacity of the condenser and, consequently, the characteristics of the circuit containing the condenser.

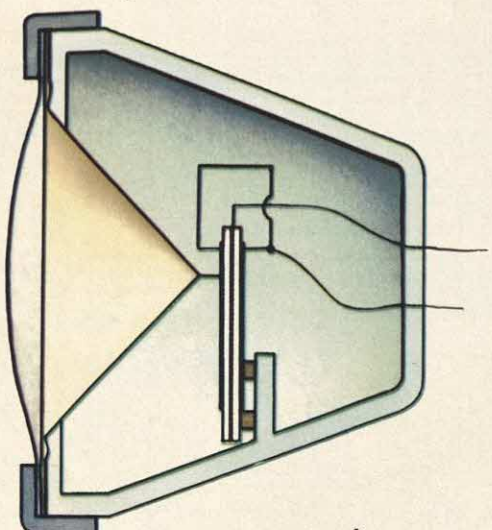
a



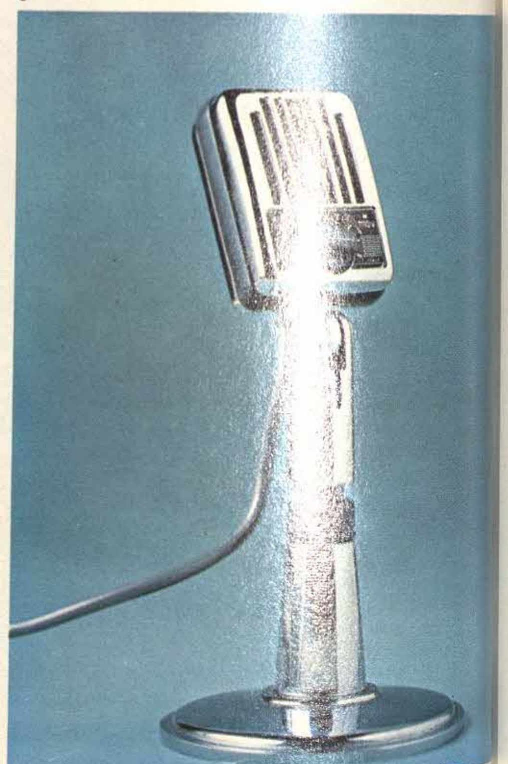
b



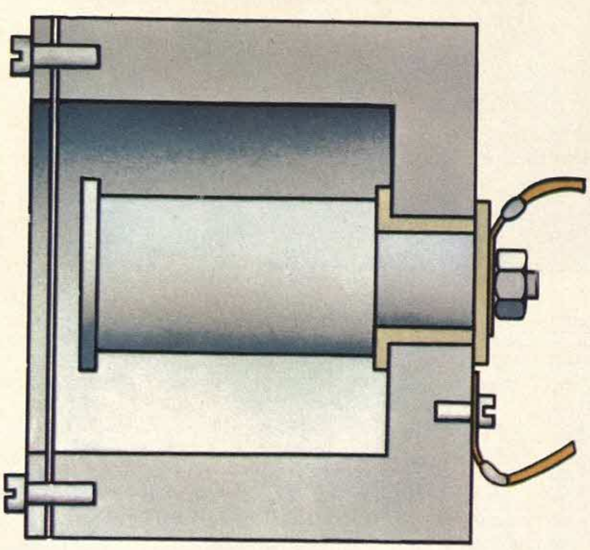
c



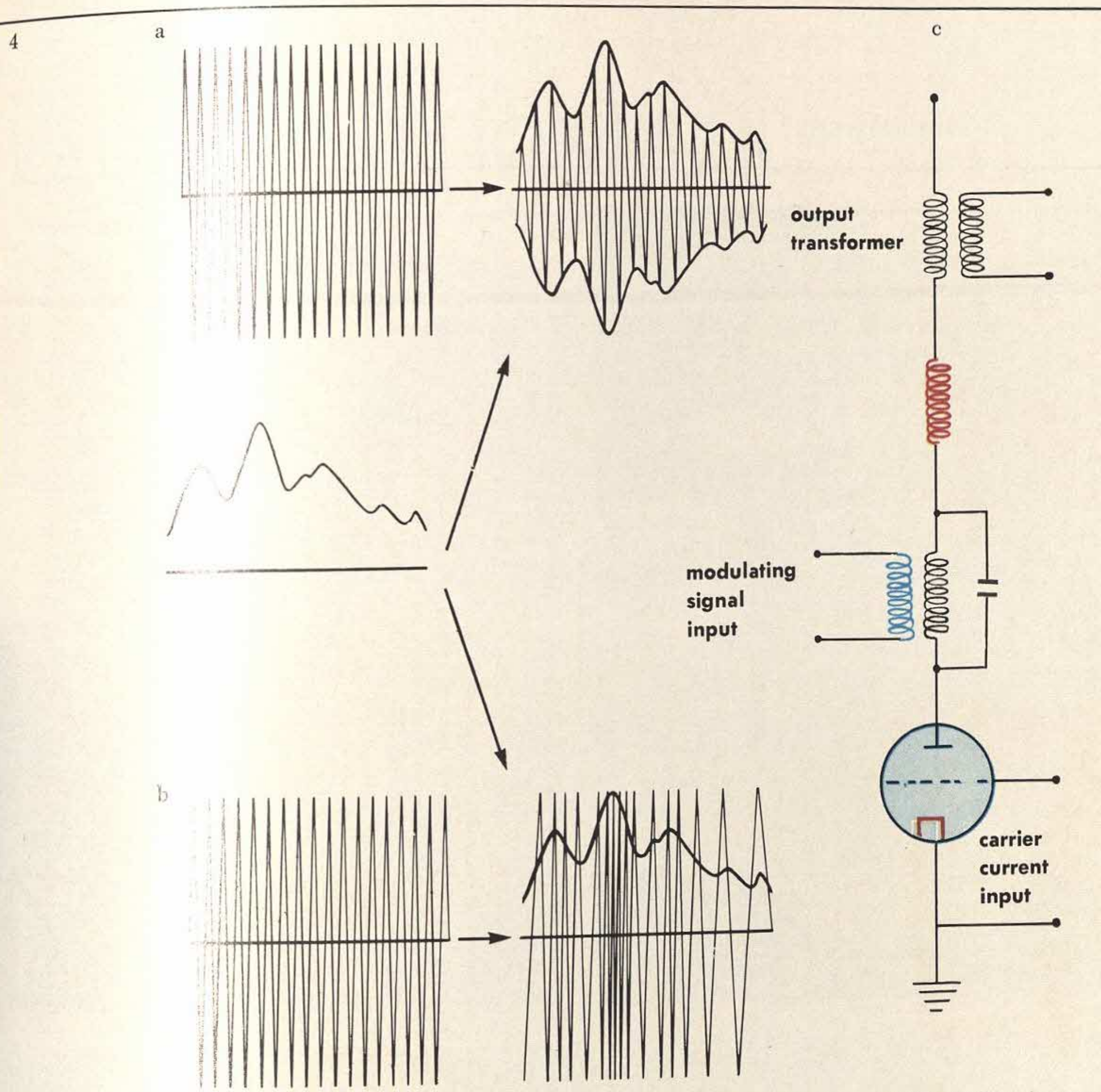
c′



d



d′



MODULATING THE SIGNAL—The low-frequency sound signal current coming from the studio must modulate the high-frequency carrier current; that is, it must transfer the signal to the high-frequency current that will then cause the emission of radio waves from the antenna. This modulation can be achieved in two principal ways. In amplitude modulation (Illustration 4a), the amplitude of the high-frequency current is varied in such a way that the shape of the waves formed by the individual sinusoidal swings reproduces the modulating current that is to be transmitted. In frequency modulation (Illustration 4b), on the other hand, the frequency of the carrier current is varied in such a way that the frequency will be high when the modulating current reaches its maximum and low when it reaches its minimum. In this way a continuous oscillation of the carrier frequency is obtained, and the amount of the frequency variation will be a measure of the amplitude of the signal to be transmitted, while the number of frequency oscillations per second corresponds to the frequency of the signal itself.

Frequency modulation requires a greater bandwidth than amplitude modulation, but it has several advantages. Among them are greater fidelity, absence of static, and smaller power requirements for a given transmission distance.

An example of the schematic layout of an amplitude modulator is shown in Illustration 4c. The high-frequency carrier current is applied to the grid of an amplifier tube whose plate circuit includes the output transformer on the antenna side. The plate circuit also includes a second transformer. The modulating signal is applied to the primary winding of this transformer, and through variations of the anode voltage, a variation is brought about in the amplitude of the modulated carrier.

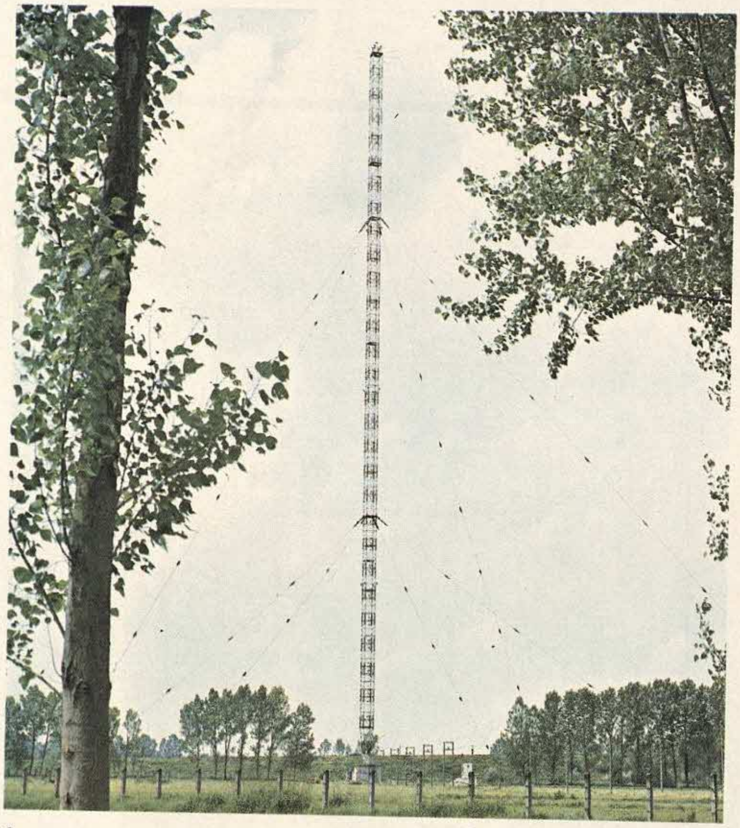
The radio waves used in radio communication range in frequency from 100 KHz (long waves) to 1,000 MHz (microwaves), and the currents from which these waves originate must oscillate with the same frequencies. The first requirement, therefore, is a series of circuits forming an oscillator whose function is to produce an oscillating current of the required frequency. This current is called the carrier current.

At the same time the carrier current is being produced, the message is being transformed into a succession of electric signals. If the message consists of sounds, these are collected by a microphone and transformed into a current whose magnitude varies with the variations in the sounds themselves. If the message consists of visual images, special photosensitive components within the television cameras generate electric signals that vary according to the brightness of the various parts of the image.

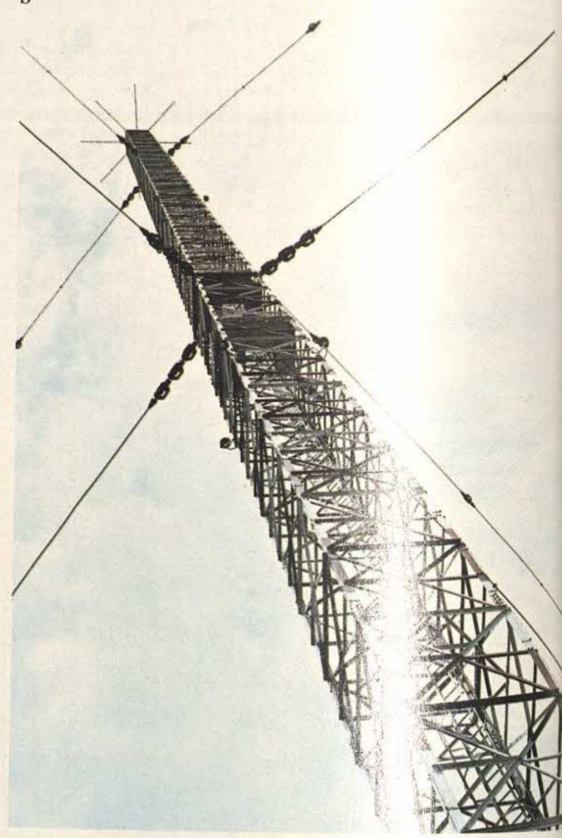
These electric signals are introduced into a circuit, called a modulator, that modifies the carrier current in such a way that it acquires the characteristics of the signals themselves. In one type of transmission, called amplitude modulation, the signal to be transmitted causes variations of the amplitude or power of the carrier current. In another type of transmission, called frequency modulation, the variations of the signal are converted into corresponding variations of the frequency of the carrier current. Whatever system of modulation is used, the final result is the same: a modulated carrier current is obtained that is first amplified and then sent to the antenna for the emission of radio waves.

5

a



b



c

d

e

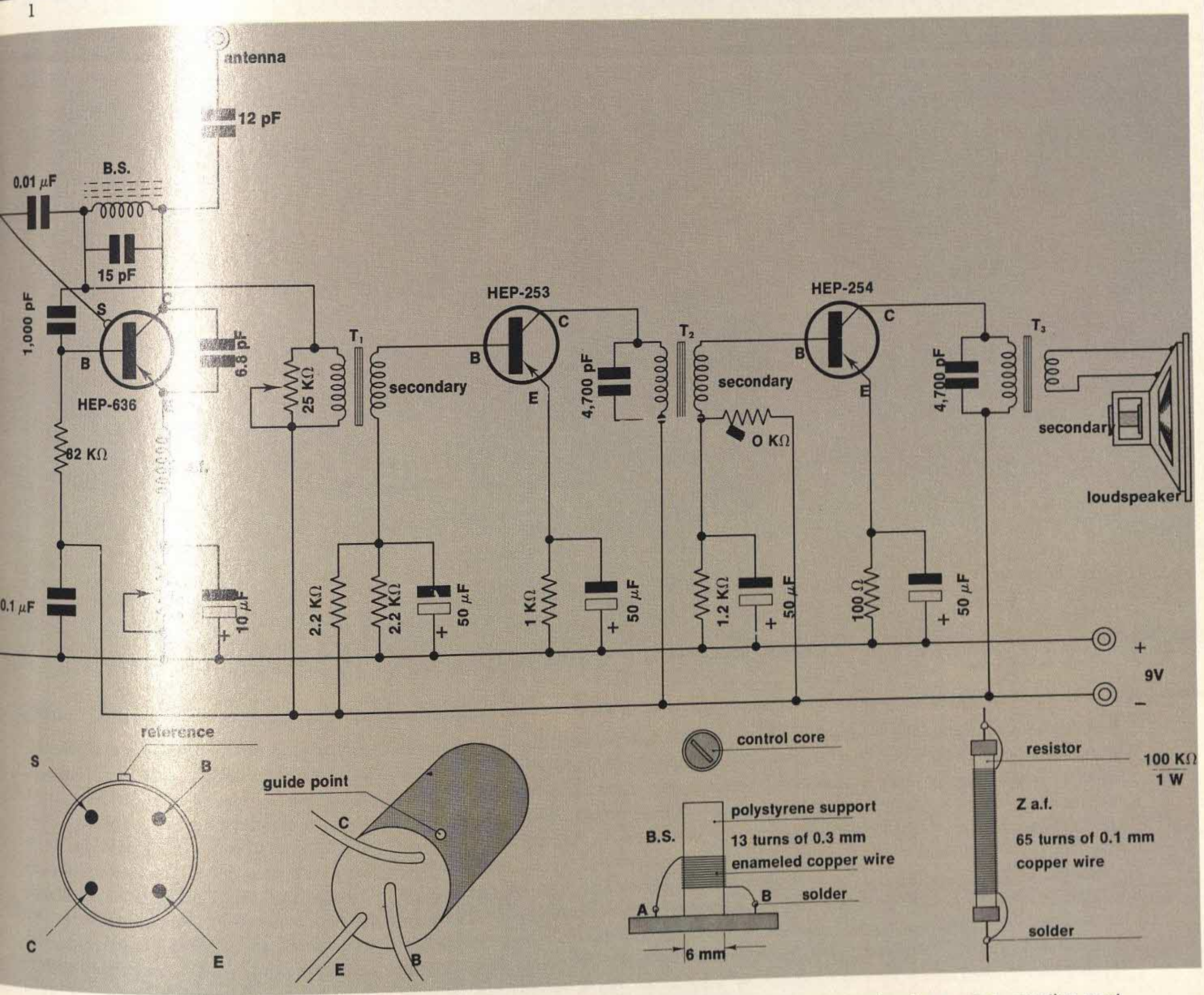
THE TRANSMISSION SYSTEM—The modulation of the high-frequency current is carried out near the antenna, which is often separate and even a long distance from the studios. Illustrations 5a and 5b show the antennas of a typical large transmitter. Illustrations 5c, 5d, and 5e are different views of a typical repeater mast.

The core of a transmission system is the high-frequency generator. This generator must produce an absolutely constant frequency in order to assure good reception and to avoid interference with nearby stations operating on similar frequencies. The stability of the frequency is assured by the use of a quartz crystal that controls the generation of oscillations. The generator is installed in a location where both temperature and humidity are carefully controlled and is mounted so that it is not affected by any mechanical vibrations in the building. Through such precautions the generator is insulated from external factors that might have a harmful effect on the constancy of its output.

The modulator circuit superimposes the signal received from the broadcasting studios, and the modulated output is then amplified. For this purpose a power of several kilowatts may be needed if the antenna has to beam the program over a large area. The electronic tubes used for the amplification are very large, often weighing as much as 50 kg (about 110 lbs). The power passing through the tubes is also great, and adequate cooling systems must be provided to dissipate the heat produced; such systems generally use water.

RADIOTELEPHONE–I | the receiver

1



CIRCUIT DIAGRAM—The receiver contains a highly sensitive detector stage formed by a Motorola HEP-636 transistor coupled to an oscillating tuning circuit, the main component of which is a coil **B.S.** with a core of magnetic material that can be moved along the coil's axis. The tuning circuit also contains a 15 pF ceramic capacitor. The core can be rotated with a plastic key or screw (metal cannot be used), which tunes the receiver to the desired station. The detector is connected to the subsequent amplification stage, a Motorola HEP-253 or equivalent transistor (note that the final stage uses a Motorola HEP-254 or equivalent transistor), by means of a transformer. In this way, a series of filters transfer the signal and also reduce the detector noise so that the signal/noise ratio is high.

The component parts of this circuit are as follows:

Stalk antenna (in common with transmitter, 1.2 m (47.2 in.) long

1	12 pF	ceramic capacitor
1	6.8 pF	ceramic capacitor
1	15pF	ceramic capacitor
1	1,000 pF	ceramic capacitor
1	0.01 μF	ceramic capacitor
1	0.1 μF	ceramic capacitor
2	4,700 pF	ceramic capacitors
1	10 μF	12 V electrolytic capacitor
4	50 μF	12 V electrolytic capacitors

In constructing the receiver, attention must be given to the polarity indicated in this illustration:

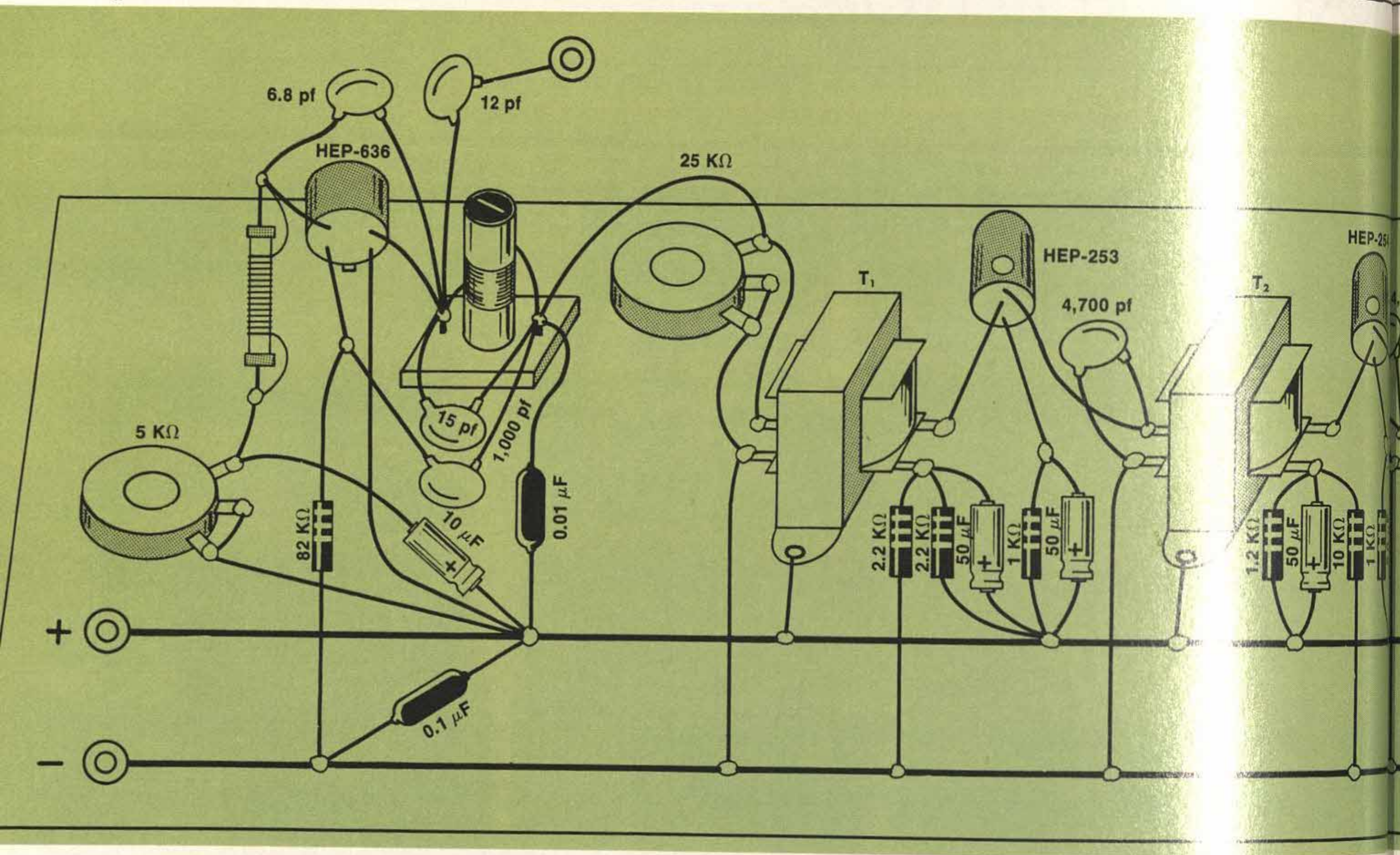
1	82 KΩ	½ W 10% resistor
2	2.2 KΩ	½ W 10% resistors
1	1,000 Ω	½ W 10% resistor
1	1.2 KΩ	½ W 10% resistor
1	10 KΩ	½ W 10% resistor
1	100 Ω	½ W 10% resistor
1	5 KΩ	miniature potentiometer
1	25 KΩ	miniature potentiometer with switch

Z a.f. inductor
B.S. coil

Transformers T_1, T_2, and T_3 are not critical. The loudspeaker is of the 16 Ω loudspeaker/microphone type.

2

a



A pair of radiotelephones permits good communications over a distance of several hundred meters (yards). Transmitters and receivers working by amplitude modulation, with a frequency of about 27 MHz, can be constructed without much difficulty.

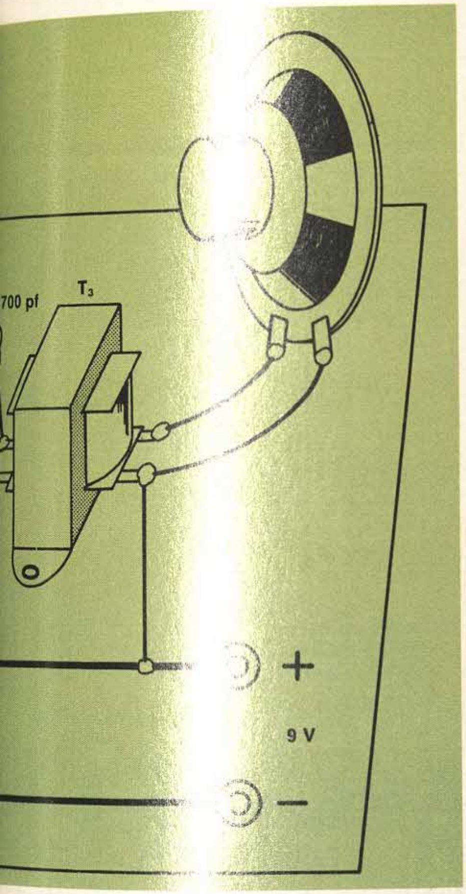
The builder of a radiotelephone set must remember that the positions of some components—and the lengths of the wires linking them—can have a great influence on the final result, even determining the success or failure of the completed instruments. To simplify this problem, this discussion is limited to construction of the receiver only; construction of the transmitter and modulator is the subject of the following article. Although separate construction of the different components is convenient for learning and understanding the radiotelephone, it is slightly more costly. In many cases, the final part of the receiver and the low-frequency amplifying stages can be used as modulators for the transmitter; the microphone used for transmission can also be used as the loudspeaker for reception. These advantages, however, involve the use of reception-transmission switches and increase the number of connections required. This more complex construction has been deliberately avoided here in order to explain simply the construction and characteristics of each component. **Before building and operating either the receiver or transmitter described in this and the following article, consult current FCC regulations.**

THE RADIOTELEPHONE RECEIVER

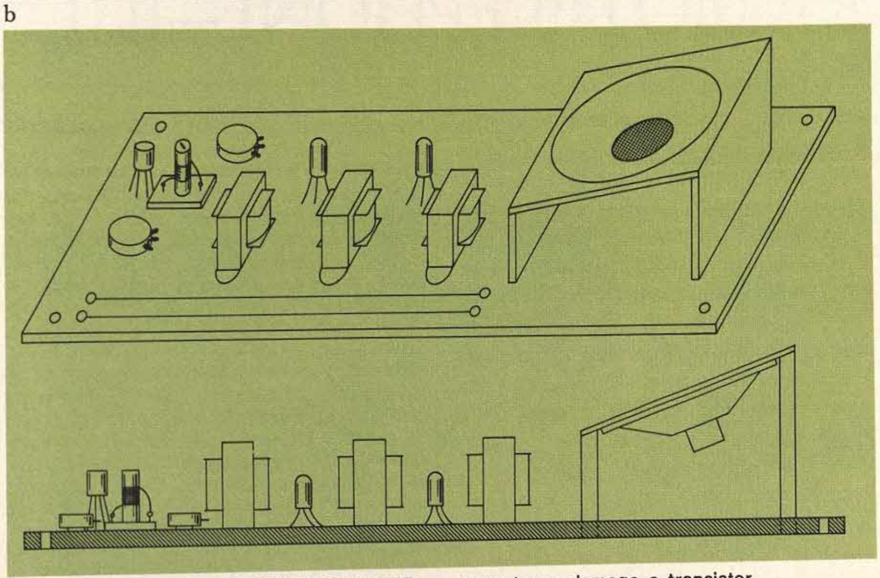
The reception of radio signals at a distance from the transmitter depends on a number of factors. Two of these factors are particularly important in radiotelephones: the power of the transmitter and the sensitivity of the receiver. Because the power of radiotelephones is limited by various factors (including legal restrictions on the power allowed for such sets), the sensitivity of the receiver has an important relationship to the radius of effective operation. This is why the circuit chosen to handle the signal received from the antenna is of the superheterodyne type.

To extract an audio-frequency (low-frequency) signal (those audible through a loudspeaker) from a modulated carrier wave (the electromagnetic wave transmitted), the carrier wave must undergo a process called detection.

Because the signal received from the receiver's antenna is normally very weak, it must be amplified so that detection of the audio-frequency signal takes place on a signal having a fairly high intensity (high amplitude). Amplification means



b



ASSEMBLY AND OPERATION OF THE SET—The layout shown in these illustrations should be followed as closely as possible. The components should be soldered with a small soldering iron, and care should be taken not to damage the contacts of the transistors with too much heat. Terminals—such as those of resistors or condensers—should not be subjected to strain by twisting or otherwise manipulating them. Resin-core solder for radio construction should be used.

When all connections have been made, the set can be checked. All connections should be checked thoroughly before testing, because a poor solder joint or the lack of a single component can damage a transistor.

Before the set is switched on the first time, all terminals joined to positive and negative poles should be checked for proper polarity. Marking these terminals red (positive) and black (negative) is an aid in making and checking these connections properly. After the two potentiometers have been adjusted to about half-power, the set can be operated. A whistle or hiss from the loudspeaker indicates that the set is working. The $\mathbf{P_2}$ (5 KΩ) potentiometer should be adjusted until the maximum noise is obtained from the loudspeaker. Volume can be adjusted with the $\mathbf{P_1}$ (25 KΩ) potentiometer.

increasing the amplitude of incoming signals. This is exemplified by visualizing a closed box containing a collection of circuits, with two wires (input) entering it on one side and two other wires (output) emerging from the other side. Suppose that a signal of amplitude 1 applied to the input corresponds to a signal of amplitude 10 emerging from the output. The amplification (gain) of such a box is 10. If the signal of amplitude 10 is fed back into the input, the resulting output would be of amplitude 100. This is the principle on which feedback circuits are constructed—in such circuits the incoming signal is amplified by being sent back to the start of the amplifying stage through appropriate connections.

This principle has operational limits. Obviously, the system cannot be used to amplify a signal infinitely. Beyond a certain critical point, parasitic auto-oscillations (noise) make further amplification impossible. Sensitivity and selectivity of a feedback receiver depend primarily on the degree of feedback; they reach a maximum when noise begins to appear in the circuit. This is the immediate vicinity of the critical point—the point of maximum sensitivity and selectivity. Operating at this maximum point is practically impossible because the circuit is extremely unstable. In superheterodyne circuits, the feedback principle can be exploited to the maximum and the drawback of instability can be avoided. The feedback is made to oscillate continuously around the critical point, at an inaudible frequency. This gains the advantages of selectivity and sensitivity without the disadvantage of instability. The single drawback of this system is that when no signal is received, the loudspeaker gives off a hissing noise; this noise, however, disappears almost completely when the transmitter signal is picked up by the receiver antenna.

Noise may be divided into two categories, man-made and natural. Man-made or industrial noise is greatest in cities and least in rural areas; its greatest cause is electrical circuit transients, which cause radiation from portions of circuits where sparks occur. Switches, motors, and ignition systems are offenders; so are x-ray apparatus, diathermy machines, industrial precipitators, high-frequency heating, and other equipment employing radio frequencies.

Noise from natural sources, entering the receiver by way of the antenna and affecting the receiver output in the same way as amplitude-modulated signals, takes two forms: atmospheric and extraterrestrial. Atmospheric noise represents the integration of field strengths from distant thunderstorms. Extraterrestrial noise stems from influences categorized as galactic, solar, and meteoric.

RADIOTELEPHONE–II | the transmitter

A radiotelephone system consists of a receiver and a transmitter; instructions for building and testing a radiotelephone receiver are contained in the preceding article. The present article is concerned with the construction of a transmitter. It is also concerned with matching the receiver with the transmitter and with testing the assembled set. **Before building and operating either the receiver or transmitter described in this and the previous article, consult current FCC regulations.**

A transmitter circuit with a quartz control oscillator is used in this example because it guarantees excellent frequency stability. This transmitter circuit has been coupled with a final amplitude modulation stage by means of a low-frequency amplifier, also called a modulator. The modulator is made with two transistors identical to the ones used in the receiver: a Motorola HEP-253 and a Motorola HEP-254 or their equivalents, with transformer coupling. The modulator transforms the weak electrical signals generated by the microphone (a loudspeaker identical to the one in the receiver) into signals that are strong enough to modulate the high-frequency signal.

The transmitter consists of two Motorola HEP-637 or equivalent transistors accompanied by their related circuits. The first transistor functions as a pilot oscillator; that is, it generates a carrier wave with a frequency of approximately 27 Mc and a high degree of frequency stability. The presence of this component in the circuit facilitates the reception of transmissions.

Radiotelephones generally function in pairs. The receiver calibrated on the first transmitter should, therefore, be mounted with the corresponding transmitter; and the receiver calibrated on the second should be mounted on the first transmitter. In this way, maximum sensitivity for each set is ensured.

The antennas must be completely extended, but must not be in contact with objects that could modify transmission (even an object that appears to be an insulator).

THE ELECTRICAL CIRCUIT DIAGRAM — The modulator section of the transmitter is very similar to the amplifying stage. The secondary of **T_4**, however, is not used; the transformer functions as an impedance, and only the primary winding is connected. The variations of current to the collector of the **HEP-254** transistor, because of signals coming from the microphone and the impedance of the **T_4** primary, cause variations in the relative voltage. This voltage, which varies in proportion to the modulating signal, is applied through **B_2** to the collector of transistor **HEP-637.** The radio frequency sent to the base of the transistor and amplified is present in **B_2,** but the amplitude depends on the voltage applied to the collector of **HEP-637** and ultimately, therefore, on the signal reaching the microphone. Briefly, the transmitter consists of two **HEP-637** transistors. The pilot oscillator (the stage that has the task of generating the radio frequency) is a quartz control, because the transmission must be extremely stable if reception is to be certain and of good quality. This stability is obtained by making a suitably shaped quartz crystal oscillate. Components of the circuit are described below.

MODULATOR

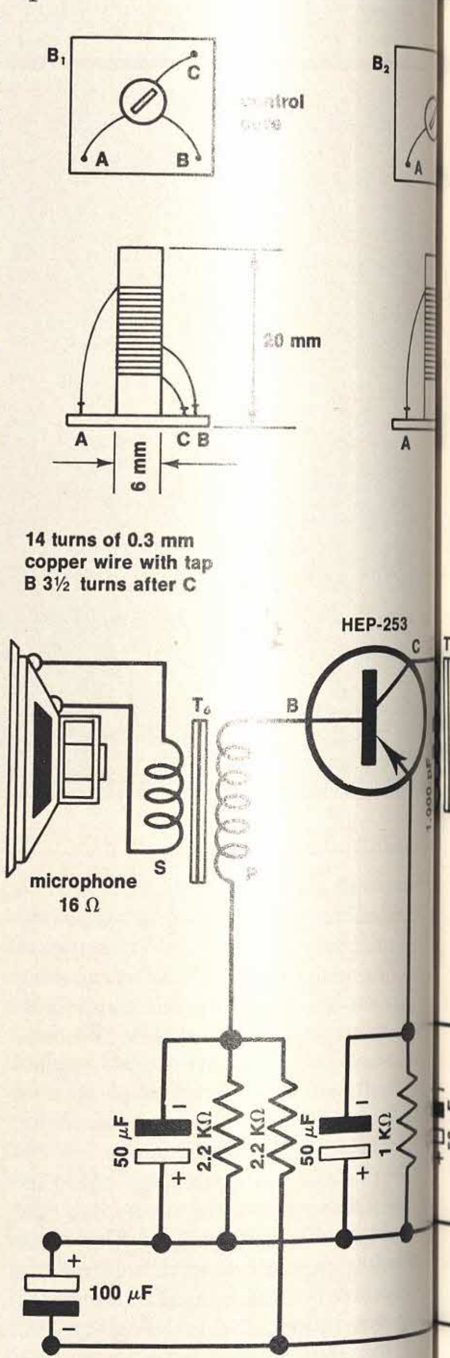
1	16 Ω microphone/loudspeaker
4	50 μF, 12 V electrolytic capacitors
2	100 μF, 12 V electrolytic capacitors
2	4,700 pF ceramic capacitors
3	2.2 KΩ, ½ W resistors
1	1 KΩ, ½ W resistor
1	12 KΩ, ½ W resistor
1	100 Ω, ½ W resistor
1	180 Ω, ½ W resistor

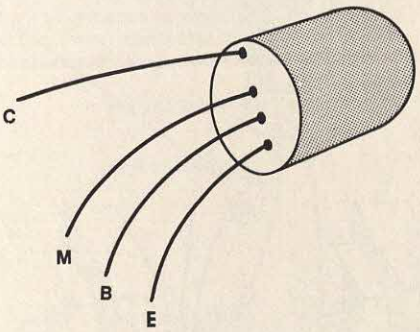
T_4, T_5, and T_6 as shown in diagram.

TRANSMITTER

3	0.01 μF ceramic capacitors
2	4,700 μF ceramic capacitors
1	350 μF ceramic capacitor
1	150 pF ceramic capacitor
1	15 pF ceramic capicator
2	2.2 KΩ, ½ W resistors
2	470 Ω, ½ W resistors
1	15 KΩ, ½ W resistor
1	18 KΩ, ½ W resistor
1	330 Ω, ½ W resistor
1	27 Mc miniature quartz oscillator
2	polystyrene bases with adjustable matching core having the dimensions shown in the diagram

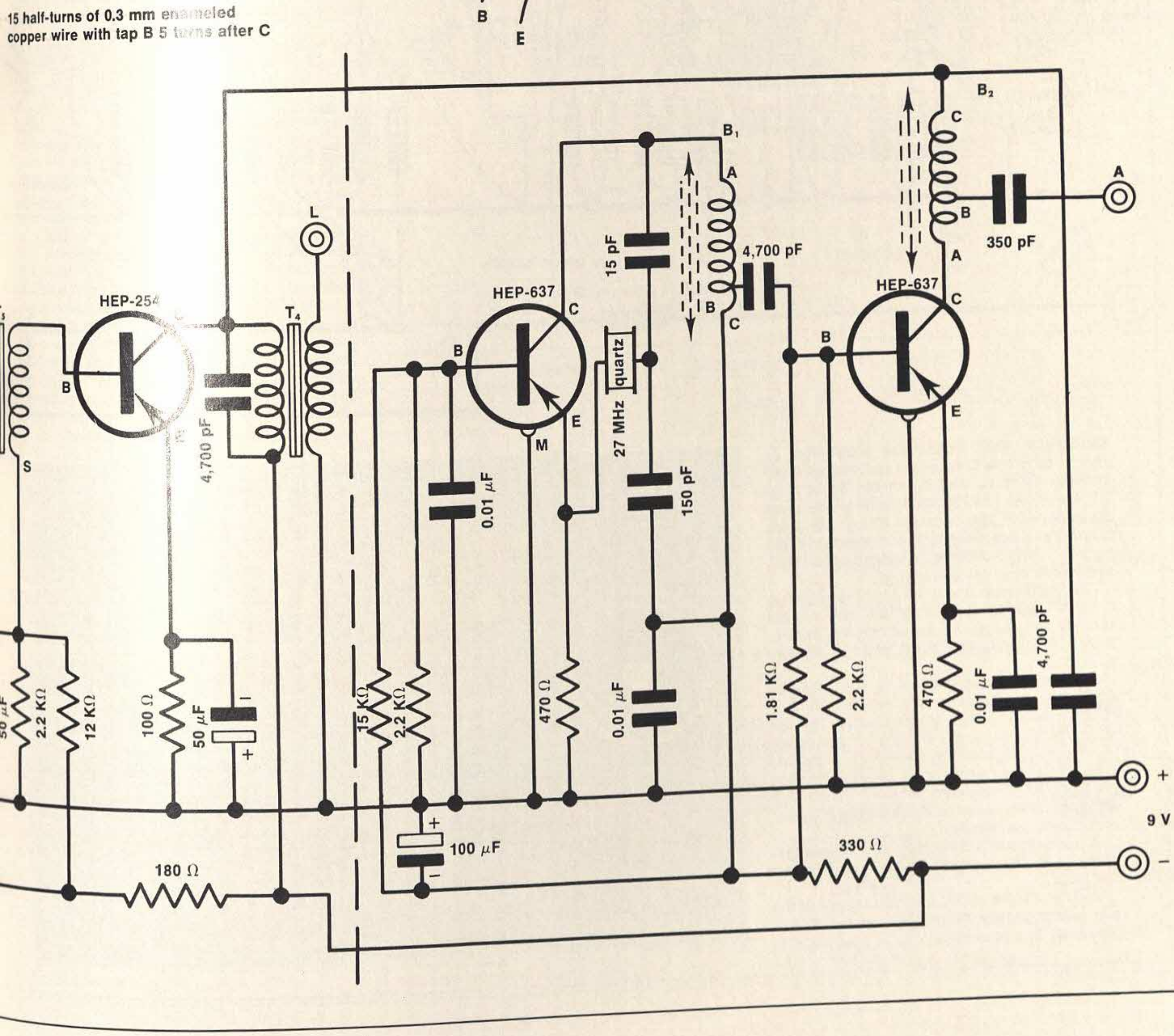
1



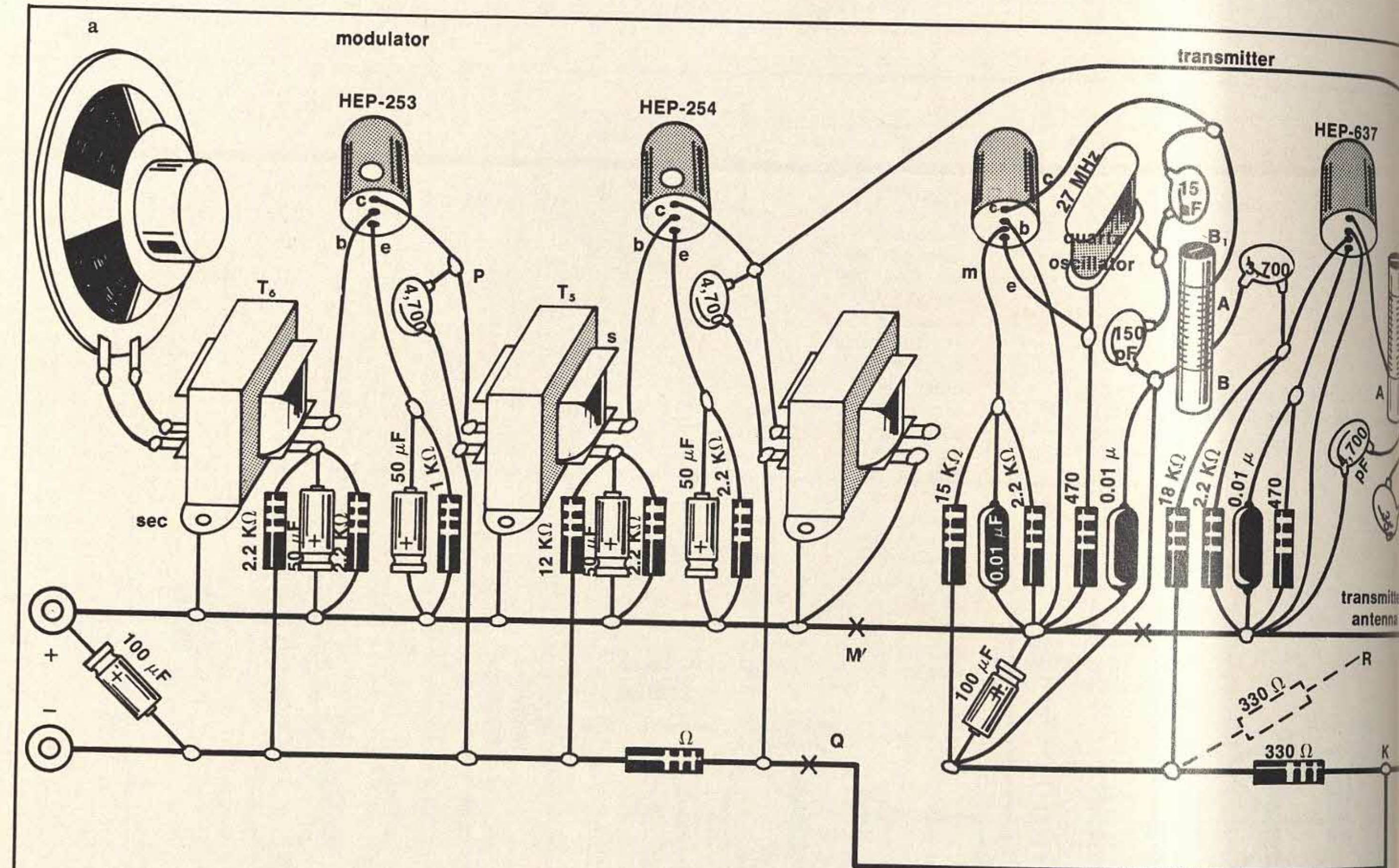


$T_6 = T_4 = HT2190$

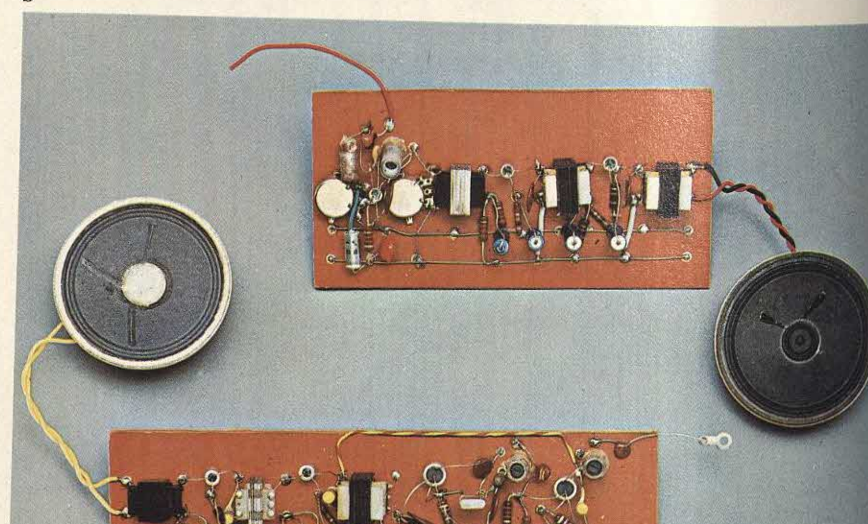
$T_5 = H333$



2



b



ASSEMBLY AND TESTING — Illustration 2a shows the optimum layout of the various components. Coils B_1 and B_2 should be at least 3.5 cm (about 1.37 in.) apart in order to avoid sparking. After the electrical circuit has been assembled according to the diagram in Illustration 1, the following calibration and test operations must be carried out.

The modulator is tested by disconnecting it at **M′** and **Q.** It is supplied with current directly by temporarily connecting the positive pole to the interruption at **M′,** and the negative pole to **Q.**

The receiver loudspeaker is connected to the transformer output with wires 1 to 2 m (3.28 to 6.56 ft) long. Words spoken into the microphone when the battery's voltage is applied to the modulator should be amplified when they are reproduced by the receiver loudspeaker. If a whistle is heard during testing, it can be eliminated by changing the position of the loudspeaker.

A milliammeter with a 15 to 30 mA (milliampere) scale range is required to test the pilot oscillator.

a. The positive circuit must be broken at the two points marked **M** and **M′** and the lead **S** removed. This temporary lead is connected to

the negative pole of the milliammeter. The positive pole should be connected to the positive pole of the 9-V battery powering the radiotelephone.

b. The terminal of the 330-Ω (ohm) resistance (as in **R**) must be detached from point **K**. This terminal must be temporarily connected to the negative terminal of the 9-V battery, with the core inserted into coil **B**. The milliammeter will then indicate a current (8 to 12 mA).

c. The core, with the previously mentioned insulated device, must be rotated and gradually pulled out of its casing. For one position, the current indicated by the instruments shows a large drop. At this point the oscillator becomes active.

The core is rotated by small fractions of a turn; as it is turned in one direction, the current increases slowly, and when it is turned in the opposite direction, the current increases sharply to the maximum value possible. The minimum should be obtained again and, turning the core away from the direction showing the maximum, the current should be set at a small amount above the minimum. The pilot oscillator is thus calibrated.

d. The 330-Ω resistance is connected to its old position at **K** and the point **N** connected to the battery's negative terminal. The terminal labeled **P** is connected to the battery's positive pole, while the milliammeter is connected as in step a. Current is introduced into the circuit from the battery and the minimum current in the oscillating circuit checked with the milliammeter. This minimum may differ from the previous one. Core B_1 can be adjusted slightly if necessary.

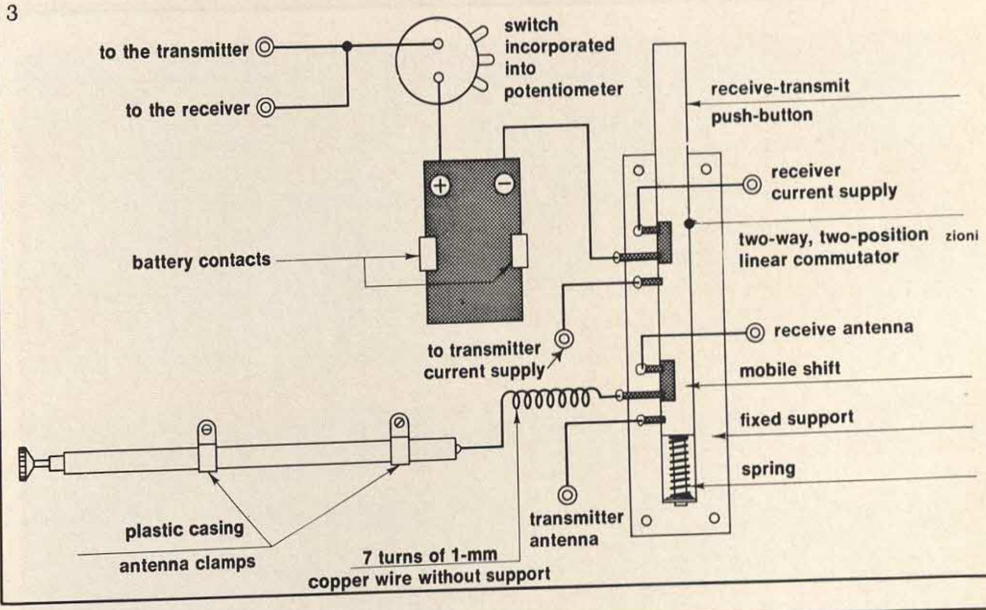
The battery is attached after this check.

e. The milliammeter is removed from its previous position and the circuit, which was interrupted to insert it, is connected at **M**. The milliammeter's negative (−) terminal is connected to point **P** of the supply line and its positive (+) terminal to the battery, with the B_2 regulation core fully inserted, as was done earlier for B_1. The antenna must also be connected to inductance **L**. The antenna must be fully extended and not in contact with, or near, metal objects. B_2 is then adjusted to obtain the minimum current registered on the instrument. After this last adjustment, the positive input is connected to the modulator at point **M'** and the milliammeter removed. The battery is connected to the circuit as planned.

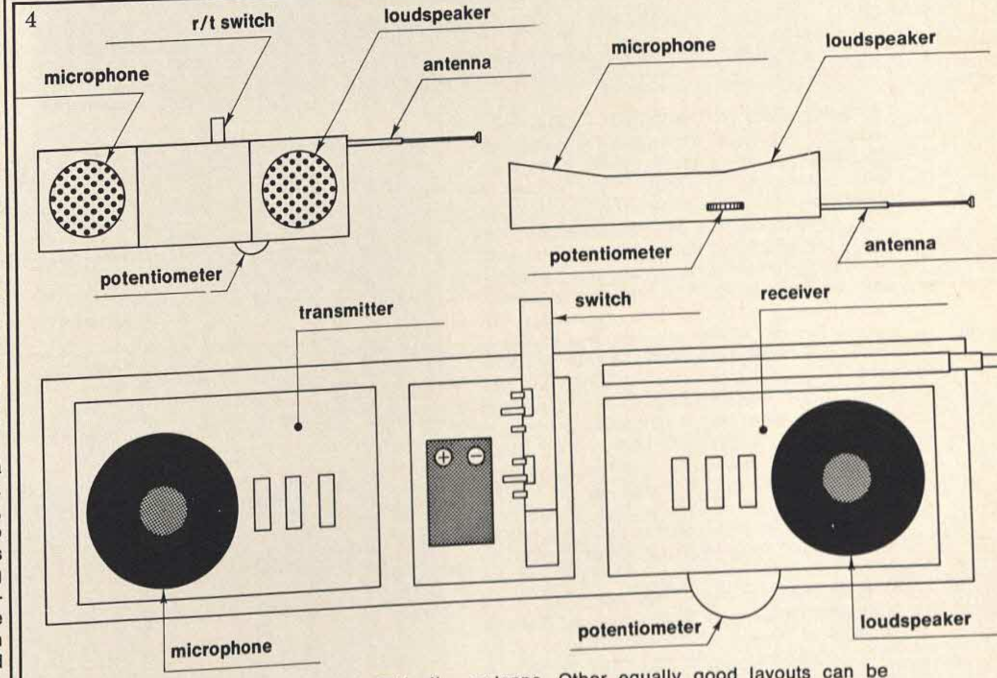
f. After the transmitter is switched on, a second battery is connected to the receiver. The receiver is not connected to an antenna, and it is placed 30 to 40 cm (about 11.8 to 15.7 in.) from the transmitter antenna. With this arrangement, the core of the receiver's tuning coil is adjusted until the transmitter is working. After this rough tuning adjustment, the receiver can be taken farther and farther from the transmitter, while the tuning is adjusted finer and finer, by means of very small rotations of the core. Illustration 2b shows the receiver and transmitter sections before connection and mounting.

RECEIVING-TRANSMITTING ANTENNA—The transmitter and the receiver use the same antenna. It must be changed from one to the other by means of a single-pole, double-throw switch. When connected in the resting position shown, the radiotelephone is receiving. The antenna is connected alternately to the transmitter and the receiver by means of an inductance by winding seven turns of 1-mm (about 0.04-in.) copper wire around a pencil.

3



4



LAYOUT FOR MOUNTED CIRCUITS—This illustration shows a possible layout for the mounted circuits, for the switch, and for the antenna. Other equally good layouts can be used, depending on how far the components and the assembly are miniaturized.

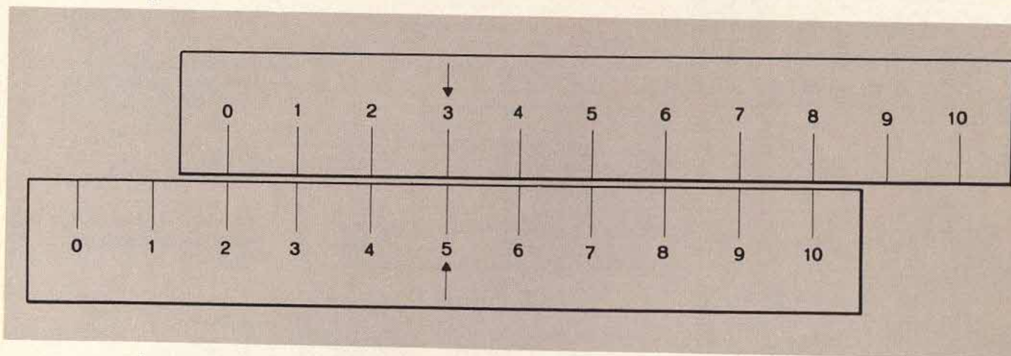
THE SLIDE RULE | the pocket computer

Many practical and technical problems—for example, working out the area of an irregular field, or the volume of the walls in a building, or the number of windings around the core of a transformer—require many arithmetic operations. If these operations were carried out by hand, they would take an enormous amount of time.

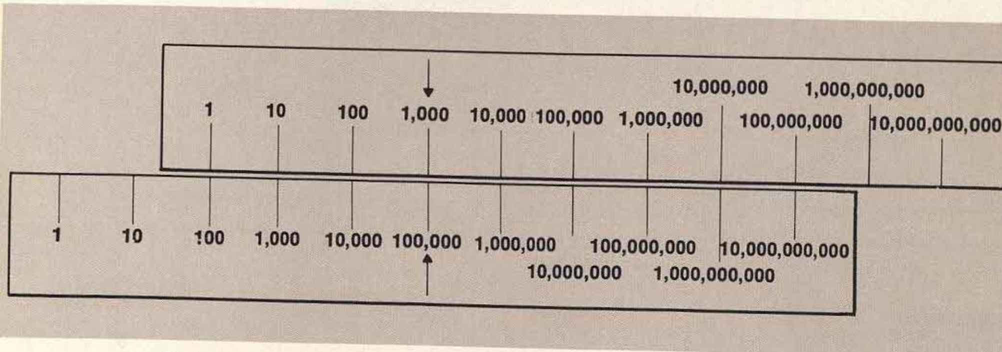
Arithmetic calculations are elementary in nature and rarely require more than knowledge of the four fundamental operations. Their length and complexity, however, introduce the risk of error, either because the probability of making a mistake increases with their length, or because fatigue makes it easier to make mistakes. To minimize error, instruments capable of performing rapidly the funda-

1

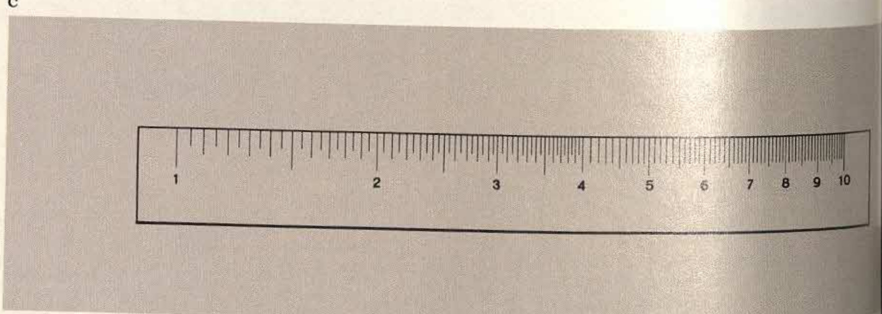
a



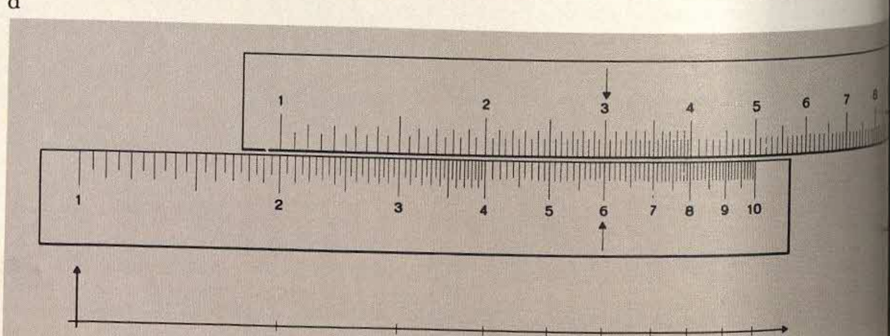
b



c



d



A SIMPLIFIED ARITHMETIC SCALE—The arithmetic scales in Illustration 1a can be used to carry out an infinite series of sums. They can also be used for subtraction, although as they stand, they cannot be used for multiplication or division. For multiplication and division, the arithmetic scales must be converted into logarithmic scales. This transformation is carried out in steps. The first step consists of simplifying the scales into centimeter divisions. The two scales are now divided into 10 parts. The scales are drawn to show the addition of 2 and 3; the sum is clearly 5. Now imagine that these subdivisions no longer represent numbers, but rather the logarithms of numbers to the base 10. Under this scheme, each division corresponds to the numbers 1, 10, 100, 1,000, and so forth, up to 10,000,000,000. The previous operation of summing 2 and 3 is now multiplication of $10^2 = 100$ by $10^3 = 1{,}000$, for which the result is $10^{2+3} = 10^5 = 100{,}000$.

Illustration 1b shows how to interpret the subdivisions of a graduated scale as the logarithms of numbers that are multiples of 10. The sum of the logarithms of two numbers is the logarithm of the product of the two numbers. As a shortcut, the number is written on the scale opposite the subdivision that corresponds to its logarithm. Thus, the scale carries the number whose logarithms are 1, 2, 3, and so forth; that is, 10, 100, 1,000, and so forth.

The slide rule is usually constructed so that it covers the numbers 1 to 10 subdivided logarithmically (Illustration 1c). This subdivision does not limit the instrument; rather, it makes it more precise in carrying out operations. Constructing the scale this way permits operations more complex than the simple multiplication of powers of 10.

The scales of a typical slide rule are subdivided as shown in Illustration 1d. The numbers correspond to their logarithms. For example, the number 2 corresponds to 0.301 parts of the total length of the scale, because 0.301 is the logarithm of 2, and so forth. The scales are drawn to shown the sum of the logarithms of 2 and 3, that is, to obtain the product of these two numbers, which is 6.

This scale yields the product in precisely the same way as the scales previously illustrated.

When it was stated that two linearly divided scales (such as on an ordinary ruler) made it possible to carry out addition operations, how to carry out the sum if it stretches beyond the right-hand end of the scale was not explained. In Illustration 1a 23 and 45 are added to result in a total of 68, a number lower than 100, the right-hand limit of the scale. How then is it possible to add 50 and 70, for example?

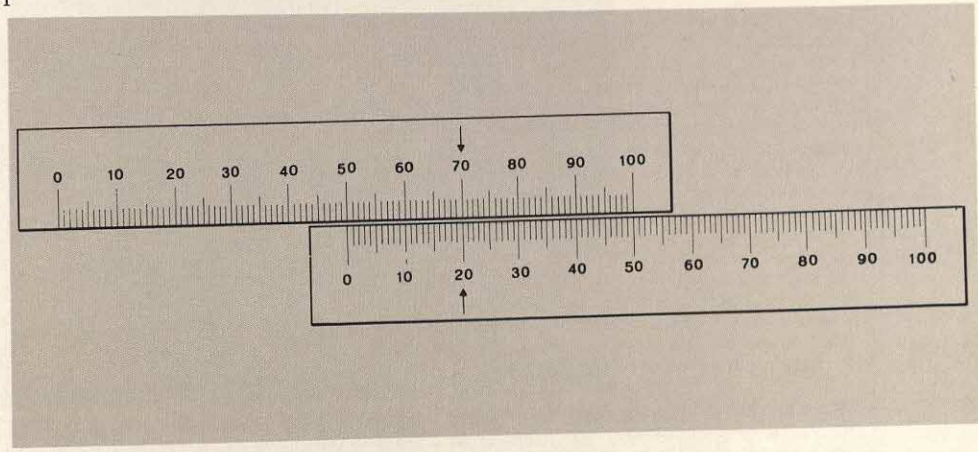
mental operations and root extraction have been invented. There are two conceptually different ways to attack the computation problem. In the first, mechanical or electronic machines carry out the required arithmetic operations in a very short time. In the second, devices that represent numbers by means of length segments are used. The machines that operate directly with numbers are called arithmetic calculating machines or digital machines (that is, a machine that works as if it were counting on its fingers). These machines are further distinguished by an adjective that specifies the mechanism used to carry out the operations (for example, "mechanical," "electrical," or "electronic" calculators).

To answer this problem, one simple observation is required. Consider two scales, one 200 mm long, and the other 100 mm long (Illustration 1e). Place the 0 of the upper scale opposite 50 on the lower scale. The 70 will then be opposite 120 on the lower scale, which consists of two identical halves covering from 0 to 100 and from 100 to 200. Now note that 100 on the upper scale corresponds to 150 on the lower scale. Since the two halves of the lower scale are identical, only half of it (the 0 to 100 portion) is needed to carry out addition operations over 100.

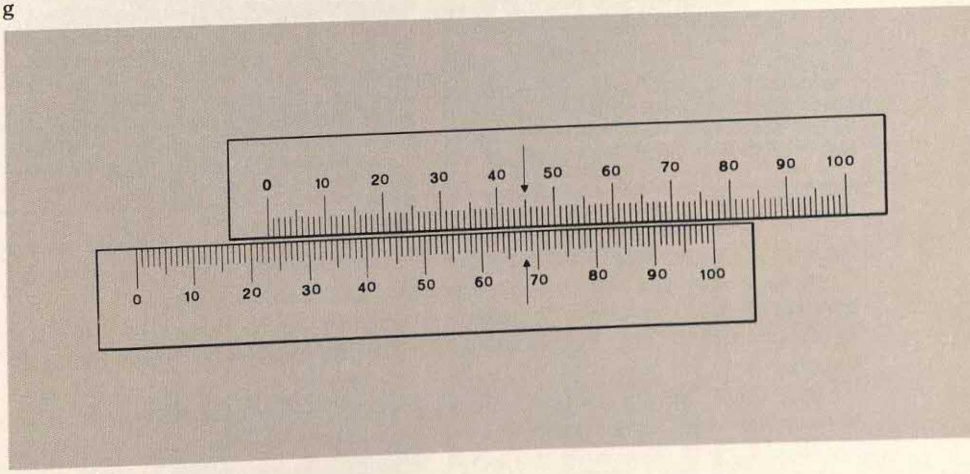
To add 50 and 70 with two scales 100 mm long (Illustration 1f), place the 100 of the upper scale opposite one of the two addends (for example, 50) and then read off at 70, the other addend, the value of the sum on the lower scale. This result is 20, but because the upper scale was moved to the left rather than to the right, it is necessary to add 100 to the result. The result is, therefore, 20 plus 100. This calculation would have been impossible if the upper scale had been moved to the right.

Subtraction is the inverse of addition, thus Illustration 1g, which shows a subtraction, is really no different from the one showing addition. Suppose 45 is to be subtracted from 68. The upper scale is moved along the lower one until the number being subtracted lines up exactly with the 68 on the lower scale. The left-hand end of the upper scale now lines up with the result of the subtraction: 23. Note that the scales are in exactly the same position they were in when 23 was added to 45. If the scales had been logarithmic rather than arithmetic, the same operations would have divided the two numbers.

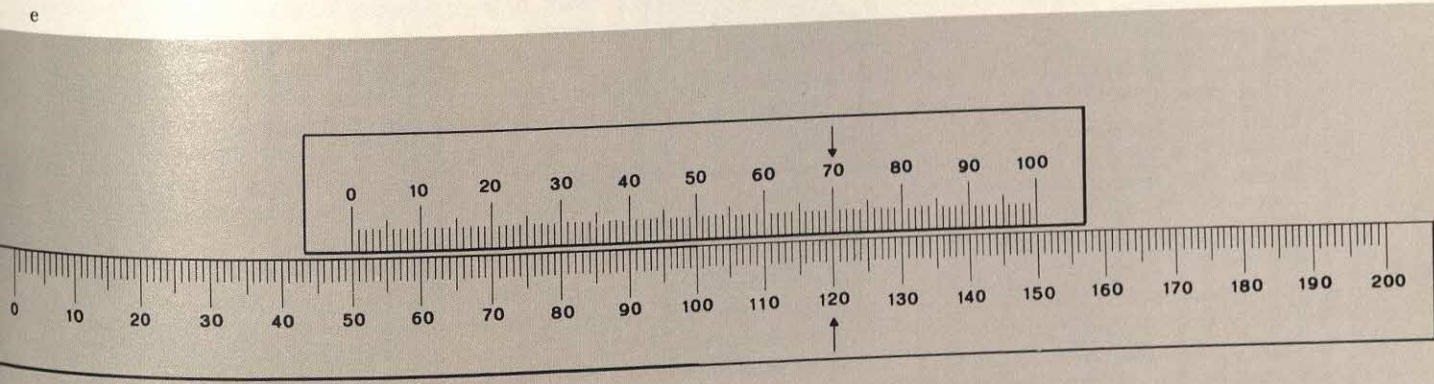
f



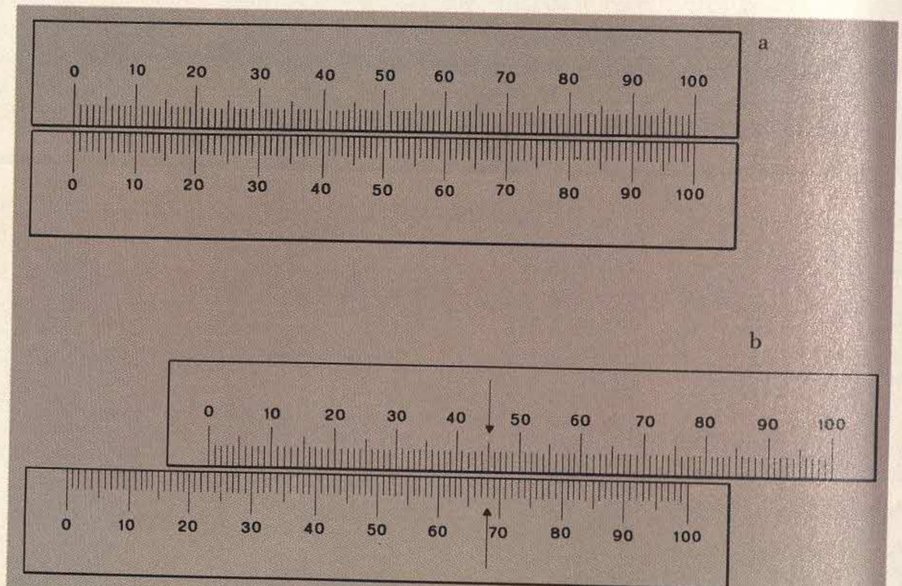
g



e



2



THE PRINCIPLE UNDERLYING THE SLIDE RULE—To understand fully the fundamental principle underlying the slide rule it is necessary to understand the concepts illustrated. A knowledge of the concept of logarithm is also required.

Suppose two wooden slats or two rectangles of cardboard are arranged beside each other so that they touch along one side. Suppose also that the two pieces are subdivided into millimeter units (Illustration 2a). Addition with these arithmetic scales is possible if a whole number is assigned to each scale subdivision. Using the subdivisions drawn for Illustration 1a, the scales represent the numbers from 1 to 100. The device now allows addition. Suppose, for example, that 23 and 45 are to be added. The principle used is as follows: if 23 and 45 represent the lengths of the segments in millimeters (mm), placing a segment 23 mm long after one 45 mm long will yield a segment as long as the sum of the two numbers—in this case, 68 mm. The operation is carried out by placing the 0 of the upper scale so that it lines up with the first addend, 23, on the lower scale. The number that corresponds on the lower scale to the 45-mm point on the upper scale is 68. To carry out this addition operation, it was necessary to move the upper scale 23 mm along the lower one.

Today's calculating machines, depending on their complexity, may carry out a very large number of arithmetic operations per second. The fastest can perform as many as a million operations per second.

Devices that carry out arithmetic operations by simulating numbers with lengths or other physical dimensions are called analog calculators.

The simplest and most common analog calculator consists of a piece of hard wood or plastic-covered wood with a groove down its center through which another piece can run. This is the common slide rule. The pieces are subdivided into appropriate units. These permit the following operations: multiplication, division, raising to a power, extracting square or cube roots, calculations with exponential or trigonometric functions, and—depending on the model—more complex operations.

Amédée Mannheim, an officer of the French artillery, invented in 1859 what may be considered the first of the modern slide rules. This rule had scales on one face only and, although it was quite simple, is basically of a type still made and designated by his name. The disposition of the scales in the Mannheim rule is the arrangement still adopted in the great majority of slide rules made in the twentieth century. This rule, which also brought into general use a cursor, or indicator, was much used in France, and after about 1880 was imported in large numbers into other countries.

In 1815 the "log-log" slide rule was invented; the fixed scale is divided into lengths proportional to the logarithm of the logarithm of the number indicated on the scale; the sliding scale is divided logarithmically.

Before 1890 slide rules were made only in England, France, and Germany, but at that time an invention led to the manufacture of slide rules in the United States. This invention introduced a revolutionary construction providing for scales on both front and back of the slide rule. An indicator with glass on both sides allowed reference to all the scales on both sides of the rule simultaneously.

Many refinements in both scale arrangements and mechanical constructions have been made since that time. The decade from 1940 to 1950 saw further developments of slide rules. Most important of these improvements was the arrangement of the scales, trigonometric and log-log, so that they operate together and at the same time maintain consistent relationship to the two basic scales. This arrangement gave added speed and flexibility to the solving of many problems, since it produced solutions by continuous operation, without the need of intermediate readings.

This article describes how the slide rule functions. Often described as the poor man's computer, the slide rule handles an enormous number of calculations met with in practical engineering and the exact sciences. Use of the slide rule is so simple and practical that anyone with a rudimentary knowledge of mathematics can master the instrument.

STEEL—I | the enormous crucible from which steel is born

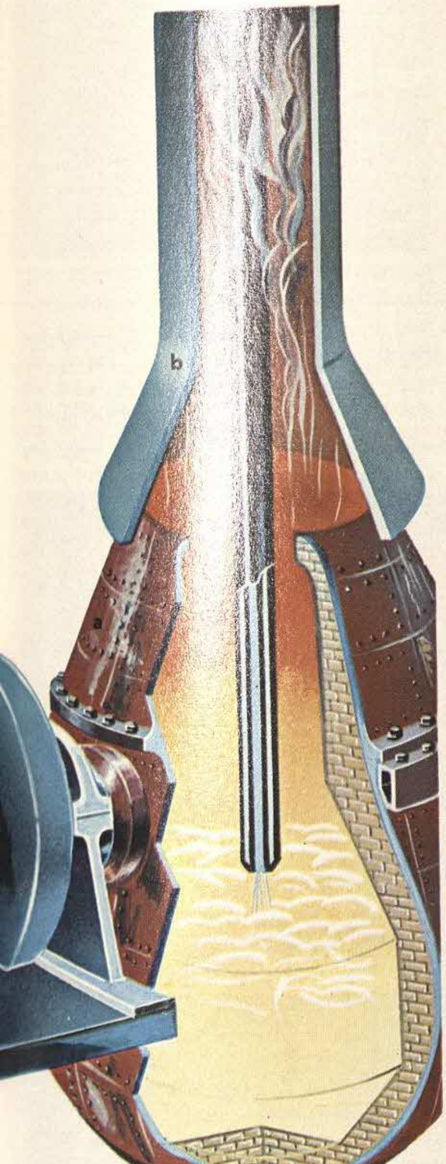
The converter process for producing steel was developed from an idea of the English inventor Sir Henry Bessemer. Bessemer knew that certain substances contained in iron (such as silicon, manganese, and carbon) combine readily with oxygen, forming oxides and producing a great amount of heat. He wished to develop a method of exploiting this principle for two reasons: to produce heat for completing the fusion of the steel bath (at the time only puddling furnaces existed for the production of steel); and to refine the iron through oxidation of the unwanted substances it contained. His solution was to pass a jet of compressed air over the liquid iron.

At first, Bessemer applied his idea to a bath of iron fused in a normal crucible. This was a container lined with refractory material, in which the metal was

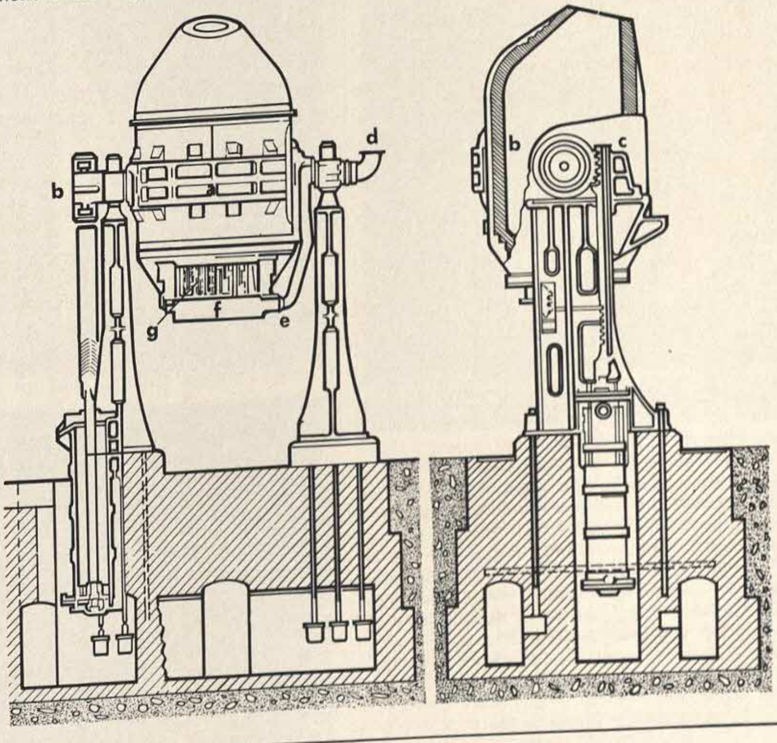
THE CONVERTER—Illustration 1b shows the converter designed by Henry Bessemer. It has the characteristic shape of a big pear with a slightly inclined neck. Inside, it is lined with refractory materials (acidic or basic) of which the central part is reinforced on the outside by a strong metal belt carrier **a** made of steel segments firmly connected by rotation pivots. One of the pivots **b** is connected by a gear to a toothed rod **c,** which is activated hydraulically and serves to tip the converter. The pivot **d** is hollow and crossed by a tube **e** carrying compressed air from a group of turboblowers. The air tube ends under the bottom of the converter in an air chamber **f** crossed by many little channels **g** in the refractory bottom, through which the air is blasted into the converter. The bottom of the converter, with the blast holes, is usually prepared with two types of refractory materials, which can be used for an average of 35 to 50 castings. The refractory materials of the walls can last for 200 castings. The Thomas and Bessemer converters rarely go over the 50-ton capacity, with a steel production that is not more than 40 tons.

The L-D oxygen converter shown in Illustration 1a is different from the Bessemer both in dimensions and in external structure. The strong metal shell **a** is completely closed at the bottom; it is supported in approximately the same way as the Bessemer, but the pivots are not hollow and they function only to support the converter. A gear system connected to a reduction gear allows for the desired rotation. The refractory lining next to the metal shell is a basic type of magnesite. Over it another lining of dolomite is in contact with the fused metal. This lining lasts for more than 200 castings. Above the mouth of the converter during operation, a hood **b** is lowered to gather the gas produced by the combustion and to convey it to a cleaner and to a boiler that recovers its heat. This hood was completely absent in the Thomas and Bessemer converters, and the gases were dispersed into the atmosphere without recovery of any of their great heat. An oxygen lance, equipped with double walls and a water-cooling system, is lowered through the hood to a height of 70 to 90 cm (28 to 35 in.) above the bath. Oxygen is blasted into the bath at a pressure of 12 atmospheres (about 175 lbs/in.²). Even the casting system is different. In the traditional converters, the load was deposited into the top and the steel was also poured out of the top after the refining. In the oxygen converter, however, the steel is cast from a small taphole at the back, a little below the mouth, thereby preventing the steel from mixing with the iron slag, which is normally cast from the top. The capacity of a modern oxygen converter can be 400 tons of steel—much greater than the capacity of older types of converters.

1a

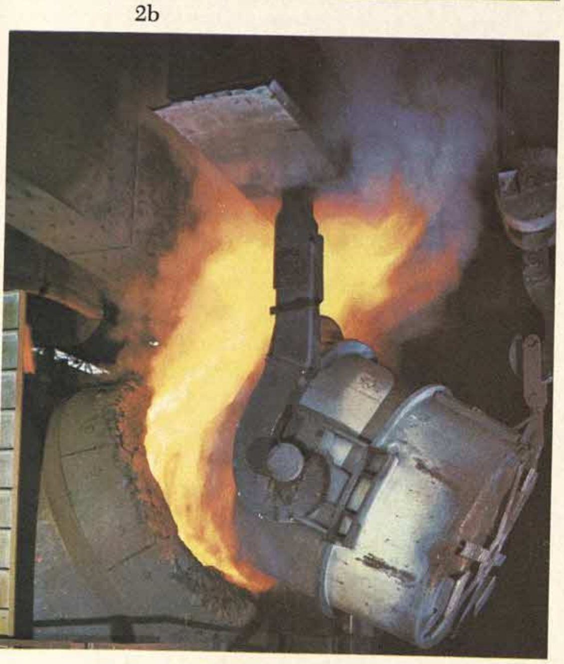


1b



2a

2b



LOADING OF THE CONVERTER—The scrap is loaded first. It is put into large containers that are lifted by a large crane and tipped inside the tilted converter (Illustration 2a). Then the converter is moved into an upright position and the slag and flux (ores and lime) descend through channels in the hood. The top is tipped again and the molten iron is poured inside. The iron comes directly from the blast furnace in small railroad cars and is poured out into a large ladle exactly as in the Martin-Siemens converter. The ladle is lifted and, with a system of levers, turned upside down over the top of the converter (Illustration 2b).

kept fused for a relatively long period. Bessemer added holes to this type of crucible and blasted air through the holes. (Air was the only available source of cheap oxygen at the time.) Later, he developed the converter, with its characteristic pear shape. This design is still in use, unmodified except for increased size over the earlier models. For a refractory lining, Bessemer used a silicon material that was known for its acidity. The Bessemer converter process is, therefore, acidic and can be used only to convert iron with a high silicon content and limited quantities of phosphorus; it is of no use for phosphoric iron.

Bessemer's process did not become widely used at first, and a certain quantity of dissolved iron oxide was allowed to remain in the final steel. This had a negative effect on its ductility and malleability. However, when a certain quantity of spiegel (spiegeleisen or spiegel iron) was added to the liquid load, mild steel could be produced in the modern way. Spiegel has a high manganese content, and manganese is an element that has a strong deoxidizing property.

In 1879, two other Englishmen, Sidney G. Thomas and Perry C. Gilchrist, decided to replace the acid lining of the Bessemer converter with a basic lining containing dolomite and magnesite. In this way, iron ore with a high phosphorus content could be used. The phosphorus in this kind of ore reacted with the oxygen, producing a great amount of heat—as silica did in the Bessemer process. The high temperature reached by the reaction of phosphorus with the oxygen initiated carbon reactions and consequent decarburization, reducing the percentage of carbon in the iron. The slag produced by the Thomas process was so rich in phosphorus anhydride that it made a good fertilizer for agricultural use.

The converter processes quickly became widely used. The Thomas process is perhaps the most common because no fuel is required and the length of time devoted to refining is short—the whole operation lasts only 20 to 25 minutes. The steel obtained is usually of a type commonly used in construction, and it can be refined even more by other processes, such as the duplex and triplex processes.

OXYGEN CONVERTERS

Experiments in mixing oxygen with the blast were tried in air-blast converters, and a more rapid and more effective reaction was produced. This enriching process, however, was limited by the resistance of the refractory liner of the furnace. These experiments showed that it is absolutely necessary to have a thermogenic element present in the iron: silicon for the Bessemer process and phosphorus for the Thomas process.

In the past few years, researchers have studied new types of converters that are slightly different from the traditional ones, even in shape. In the new converters, pure oxygen is blasted directly over the bath. These oxygen converters convert iron to steel much faster than their predecessors; moreover, they impose no strict limitation on the composition of the iron that is to be used. The use of oxygen permits the required temperatures to be reached even if large quantities of thermogenic elements are not present (large quantities of these elements cause an excessive production of slag). Considerable savings can also be effected with an oxygen converter by using a certain amount of scrap iron for loading the converter.

Of these methods, the L-D (or Linz-Donawitz) oxygen process is particularly widespread. The process, developed in Austria in 1952, allows for direct transformation of liquid iron into steel with a final addition of 30 to 35 percent scrap iron. This system uses a simple blast, at ultrasonic speeds, against the surface of the fusion bath. A jet of very pure oxygen comes down from the top by means of a water-cooled lance. An improved version is widely used, especially in the United States, because of the high-quality steel that it produces. Highly automating the process would probably enable the iron and steel industry to reach qualitative and quantitative levels heretofore unknown. This very modern process, destined in a very short time to become preeminent over all other systems, merits special attention.

3

SPEED OF THE CONVERTER—Following loading, the converter is returned to its vertical position. The hood, visible in the photograph, is lowered and blasting with the oxygen lance begins. This operation is precisely and completely controlled electronically from a nearby location. All reactions that occur inside the converter are similar to those that occur in the Martin furnace, but they occur with much greater speed and force. The temperature immediately rises because of the effect of the oxygen; decarburization begins in 5 or 6 minutes, from the oxidation reactions to the thermogenic elements in the iron. A large production of carbon monoxide characteristically colors the flame. During this phase, muffled noises caused by the boiling of the fused mass in reaction to the carbon monoxide can be heard. This is truly a miniature eruption that can be controlled from the outside. Observation of the flame (first yellow, because of the sodium, and then blue because of the carbon) indicates how the operation is proceeding. When this phase is finished and the material loses its color, decarburization is just about finished.

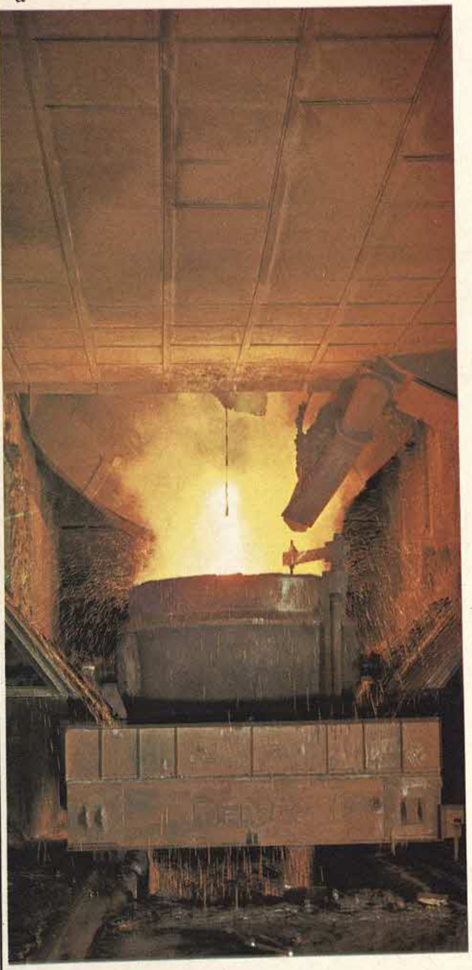
4

TEMPERATURE—At [illegible] end of the refining process, the converte[illegible]ipped again and the end of the long tube [illegible]roduced into the top (Illustration 4a). This [illegible] is part of a pyrometer, which measure[illegible]xtremely high temperatures. It measures th[illegible]mperature of the bath, which reaches 1,800° C (almost 3,300° F). After other slags and alloy metals have been added (blocks of aluminum may be seen at the right in Illustration 4b), the bath is left to settle, to provide time for the steel composition to homogenize.

a

b

a



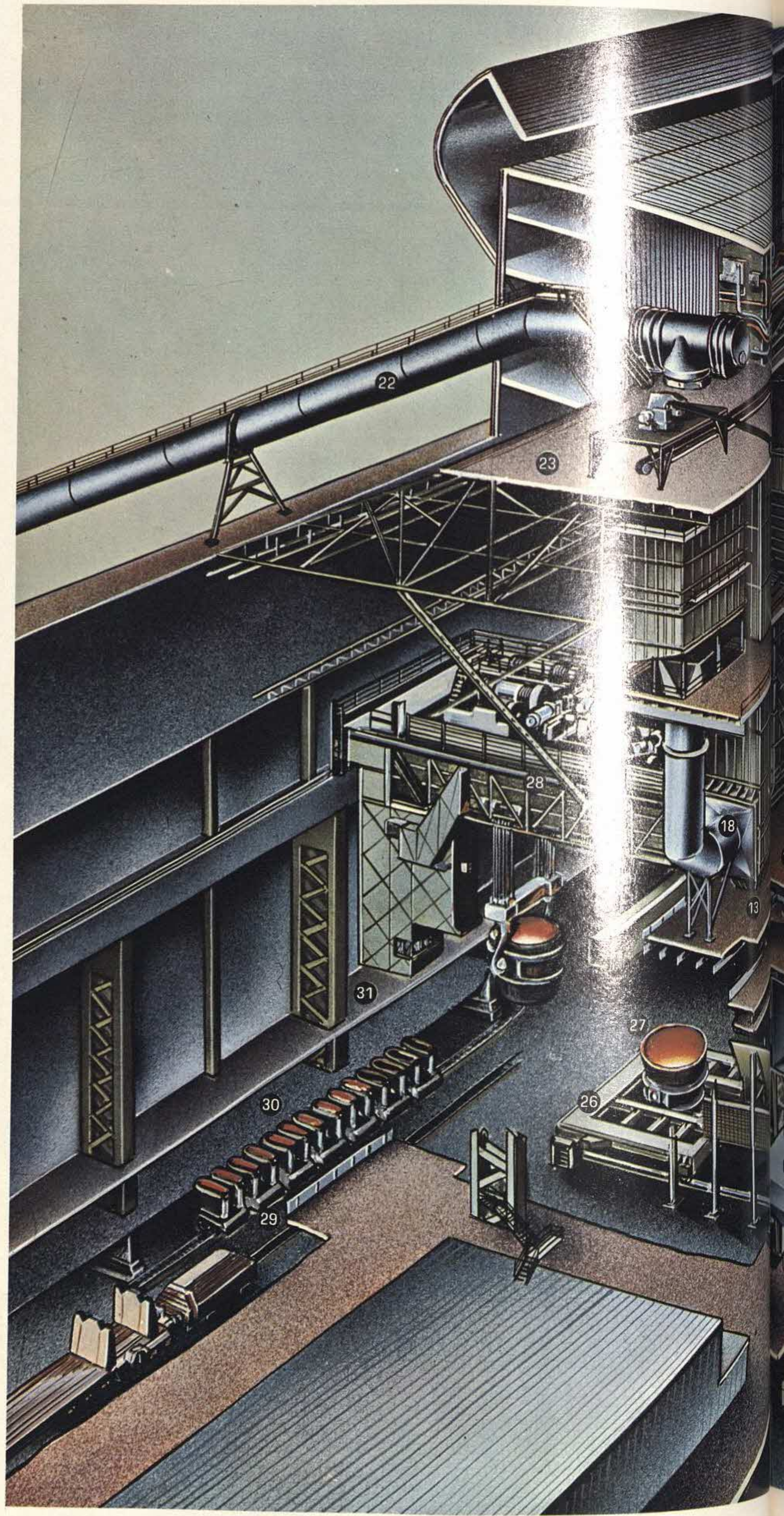
b

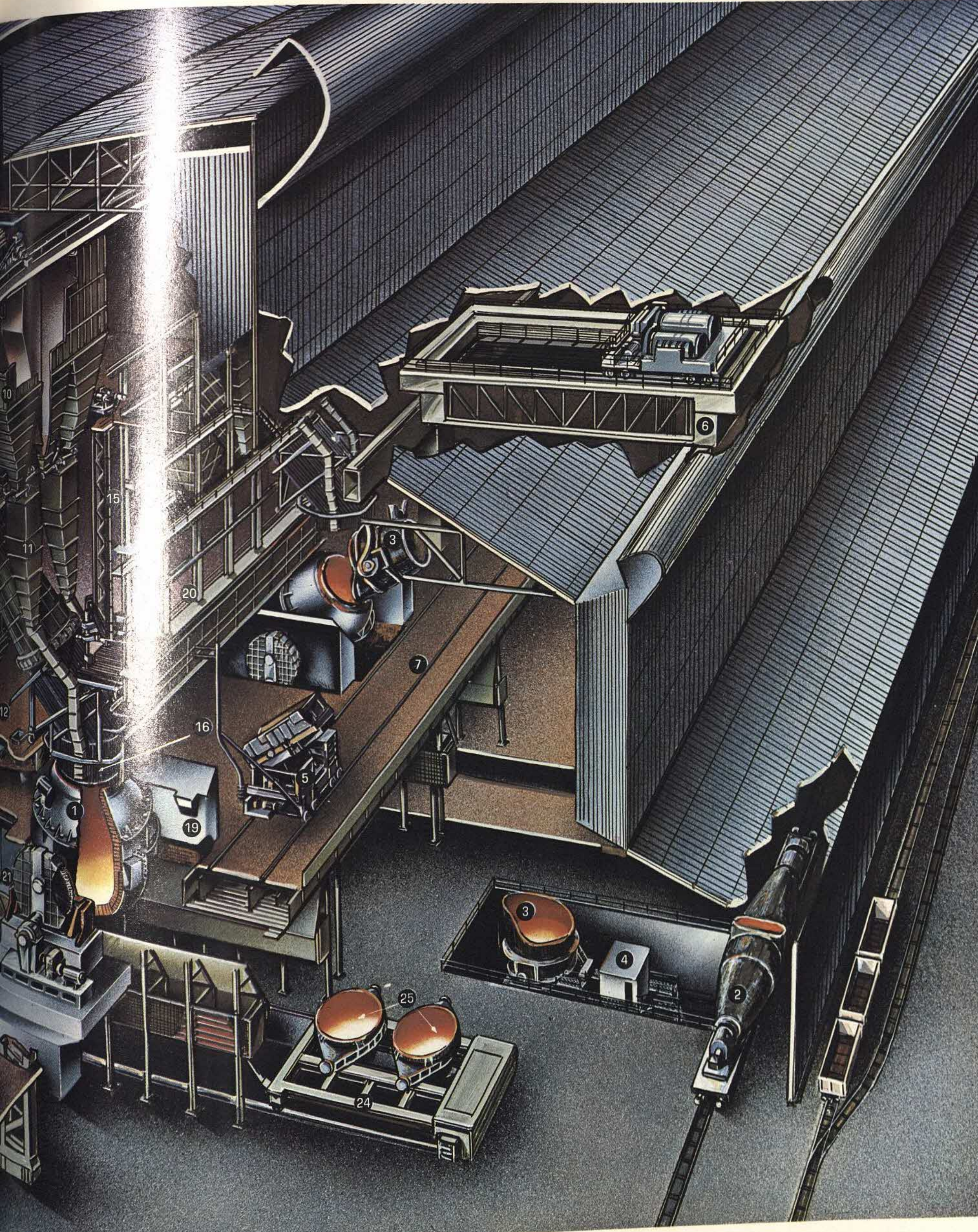
5

CASTING—Casting is done into a ladle as for the Martin furnace. The ladle is placed on a carrier (visible in Illustration 5a) that is pushed underneath the converter. The converter is tipped slowly and the steel flows out of the casting tap. After the steel has been cast, the converter is tipped in the opposite direction and the slag is poured out. Other slags that deoxidize the bath are thrown into the ladle (Illustration 5b). The slag floats on top of the steel and keeps its surface from cooling excessively, thus preventing useless waste of metal and difficulties during the operation of the casting into ingots.

AN ILLUSTRATED SUMMARY OF AN L-D STEEL MILL

1. The converter
2. Rail car
3. Iron ladle
4. Cabin for weighing the liquid iron
5. Car loaded with scrap
6. 400-ton loading crane
7. Loading level
8. Belt for incoming ores
9. Belt for distributing ores
10. Ore hopper
11. Weighing hopper
12. Additive chute
13. Servicing level
14. Oxygen lance
15. Lance truck
16. Mobile hood
17. Boilers for recovering heat
18. Smoke purifier
19. Control cabin for converter
20. Flexible tubes for oxygen and cooling water
21. Group of reduction gears for rotation of converter
22. Smoke pipe
23. Bin level
24. Cars for transferring iron caldrons
25. Iron caldrons
26. Car for transferring steel ladle
27. Steel ladle
28. Crane for 400-ton casting
29. Rail cars for ingot molds
30. Ingot molds
31. Casting platform




10
15
11
20
12
16
1
19
21
3
7
5
6
3
4
2
25
24

STEEL—II | the hot-rolling process

The production of steel, whether by the open hearth process, or by the converter process, results in cast ingots. With few exceptions, these ingots are of the ordinary type—constructional steel—and have different mechanical characteristics (and different chemical compositions), according to the uses for which they are intended.

The ingot is the starting point in a series of working processes that lead to many different products: sheets, girders, bars of various forms and thicknesses, rails, rods, strips, and sections of widely differing shapes. Whatever the final product, the working process, called "rolling," is carried out in several stages. The machinery through which the ingot or the intermediate product is forced to pass in order to change its form is called a "rolling mill." A rolling mill consists of two exceptionally strong rollers of cast iron or steel mounted on horizontal and parallel axes and rotated in opposite directions inside a solid, fixed framework (called the housing). The incandescent ingots are inserted between the two rollers, where they are moved along by friction until they gradually assume the desired shape.

Leonardo da Vinci, the versatile Italian artist, scientist, and inventor, was the first person to visualize the use of this means of carrying out different operations on various metals. What is more, he also discovered the principle (which found practical application only within the last half-century) of using four rollers, the two outermost of which perform only the function of resisting the force required to roll the ingot. Leonardo employed rollers made of hard bronze mounted on an iron core and used them for rolling tin blocks into sheets. The first operations of this kind were generally carried out on metals that could be deformed easily even in the cold state—tin and lead, for example. Subsequently, the process was gradually improved, and the smooth rollers were replaced by shaped rollers for different forms. The development of iron metallurgy, leading to the production of large quantities of steel, eventually gave further impetus to numerous and varied improvements, resulting in the extensive present development level of this branch of metallurgy.

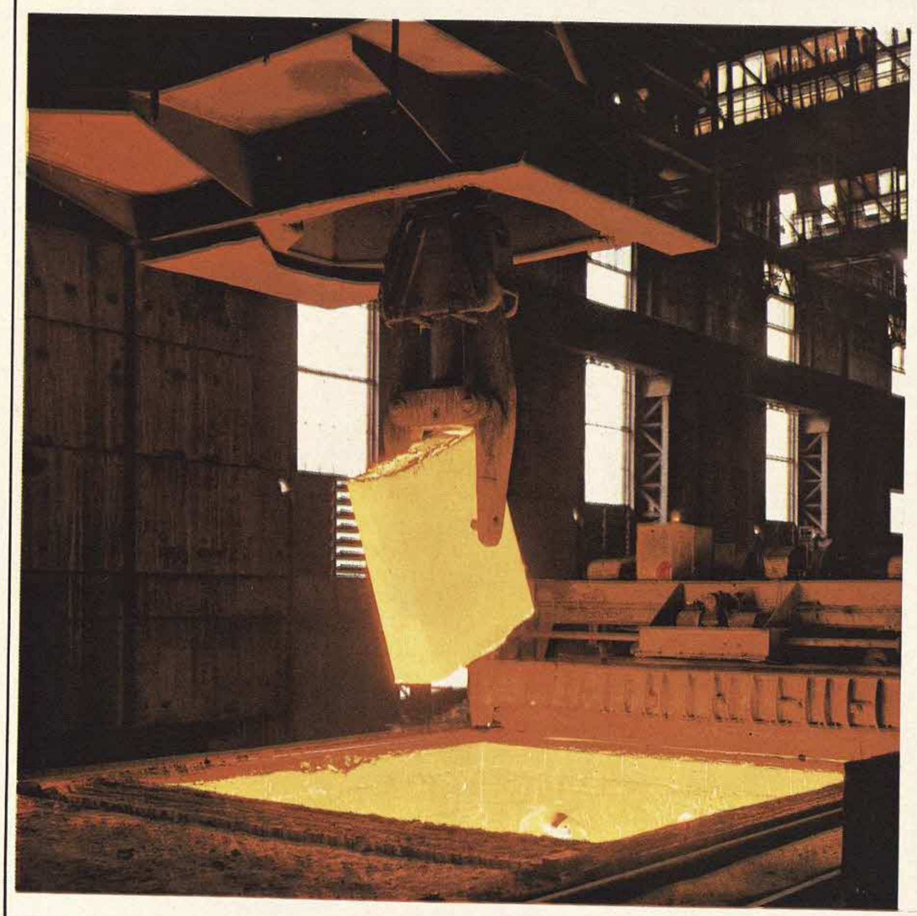
HOT-ROLLING OF STEEL

Rolling processes are of two types: hot-rolling and cold-rolling. Hot-rolling generally uses ingots as the starting material and produces a series of forms that can be utilized readily in mechanical and structural construction. Examples are sheet steel of various thicknesses in rolls or plates, beams, rails, round bars, and similar forms. In cold-rolling, on the other hand, these intermediate products serve as the starting materials from which thinner products are obtained; they are also used to cover surfaces with protective facings. Zinced (galvanized) or tinned sheets are obtained in this way.

The specific process examined in this article is hot-rolling. In the hot-rolling process, steel is poured into the ingot molds and the mass of metal begins to solidify. During solidification, however, an ingot may be subject to modifications

1

a



PIT FURNACES—Pit furnaces consist of cells of various sizes, with walls lined with refractory materials based on silica or magnesite. This type of furnace is generally heated by mixed gases—a mixture of blast-furnace gas and coke-oven gas. Naphtha and methane can also be used as fuel. The gases, after being preheated in heat-regeneration chambers similar to those used in the open hearth furnace, enter the pit by means of burners placed at the sides of the furnace, as shown in Illustra-

b

that could cause defects to form in the mass of the metal: for example, the metal may be subjected to very high internal stress or to chemical reactions. If the potential defects are not successfully prevented, they can cause considerable damage by weakening the final product. The following defects and their causes are of major concern in steel production:

1. *Pipes*—Cavities formed as a result of shrinkage are called pipes. They are formed during the solidification process in places where the metal has remained in the liquid state for the longest time. The walls of pipes do not weld together, and this defect must be eliminated in advance. The formation of pipes can be prevented by casting the ingot by the method of "uphill teeming."

2. *Seams*—The defects known as seams result from steel spray at the bottom of the ingot mold. Small seams are eliminated during the reheating of the ingot prior to hot-rolling.

3. *Cracks*—Fissures or cracks in the surface of the ingot are caused by uneven cooling or by imperfections in the walls of the ingot mold. They are eliminated during the rolling process through the use of either an oxyacetylene torch or a chisel.

4. *Blowholes*—Considerable quantities of gas (carbon dioxide, nitrogen, hydrogen) are found in the mass of fused steel. The solubility of these gases diminishes during solidification, leading to the formation of bubbles, which rise to the surface for as long as the fluidity of the metal permits. When the steel is cooled to a certain point, the gas bubbles can no longer rise; then they remain imprisoned in the metal as so-called blowholes. The blowholes tend to disappear when the ingot is reheated prior to rolling.

The foregoing review of defects that can occur in ingots emphasizes the fact that the reheating of these ingots prior to rolling not only renders the metal softer (and thus more readily workable) but also performs the important task of eliminating most of the solidification defects. The reheating is generally carried out in pit furnaces consisting of many vertical cells, into which the ingots are loaded from above. When the temperature of an ingot has become stabilized (the time required for stabilization depends on the size of the ingot), the ingot is removed from the furnace and sent to the rolling mill. Ingots leave the rolling mill transformed into various shapes, according to the processing that they are to undergo. Flat blooms, weighing several tons each, are obtained from very large ingots and are intended primarily for the production of sheet steel. Blooms of much smaller dimensions and weights are obtained when the intermediate product is intended for the manufacture of beams, rails, strips, and similar products. Billets several meters long with square, relatively small cross sections are produced when the steel is intended for conversion into round bars or hot-drawn wires. Examination of the accompanying photographs will help to provide an appreciation of the meanings of these various terms. The illustrations concentrate on flat blooms, following them on their way through the steps involved in turning out a quality product.

tion 1a. When the ingots have been heated to the temperature required for rolling, they are picked up by a large, overhead, traveling crane equipped with a double-jawed lifting device. This device must be immersed in water frequently to cool it and to prevent breakage. The traveling crane deposits the hot ingot on a special trolley (Illustration 1b), which takes it to the rollers of the rolling mill train. A battery of furnaces can also be seen in this photograph.

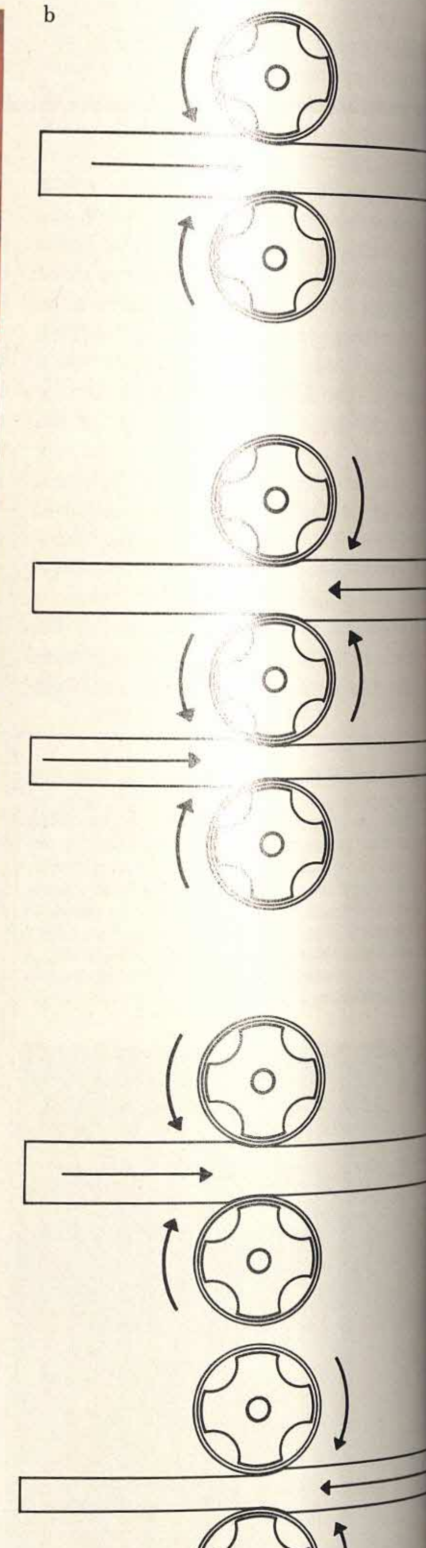
2

WEIGHING THE INGOT—The transport trolley discharges the steel ingot onto a roller conveyor, which transports it to the rolling mill proper. Immediately prior to this, however, the ingot passes over a weighing device operated from the control cabin; the device measures the gross weight of the ingot. This enables production-control personnel to determine the exact quantity of steel being processed and, by comparing the gross weight of the ingot with the weight of the finished products, to obtain a measure of the losses incurred during rolling (trimmings, defective parts, scale). The material lost in these ways can generally be recovered and reprocessed in the form of scrap steel.

BLOOMING MILL — Illustration 3a shows the "roughing" mill, which transforms the ingots into blooms, flat blooms, and other forms. Characteristic features of a rolling mill are the housings that contain the rollers; these rollers, in turn, form a "train." The rollers may be arranged in different ways (Illustration 3b); twin trains, triple trains, and double-twin trains are possible arrangements. The roughing train is generally a twin train of the reversible type—that is, a train that can pass the ingot first in one direction and then in the other, as the motor

a

b



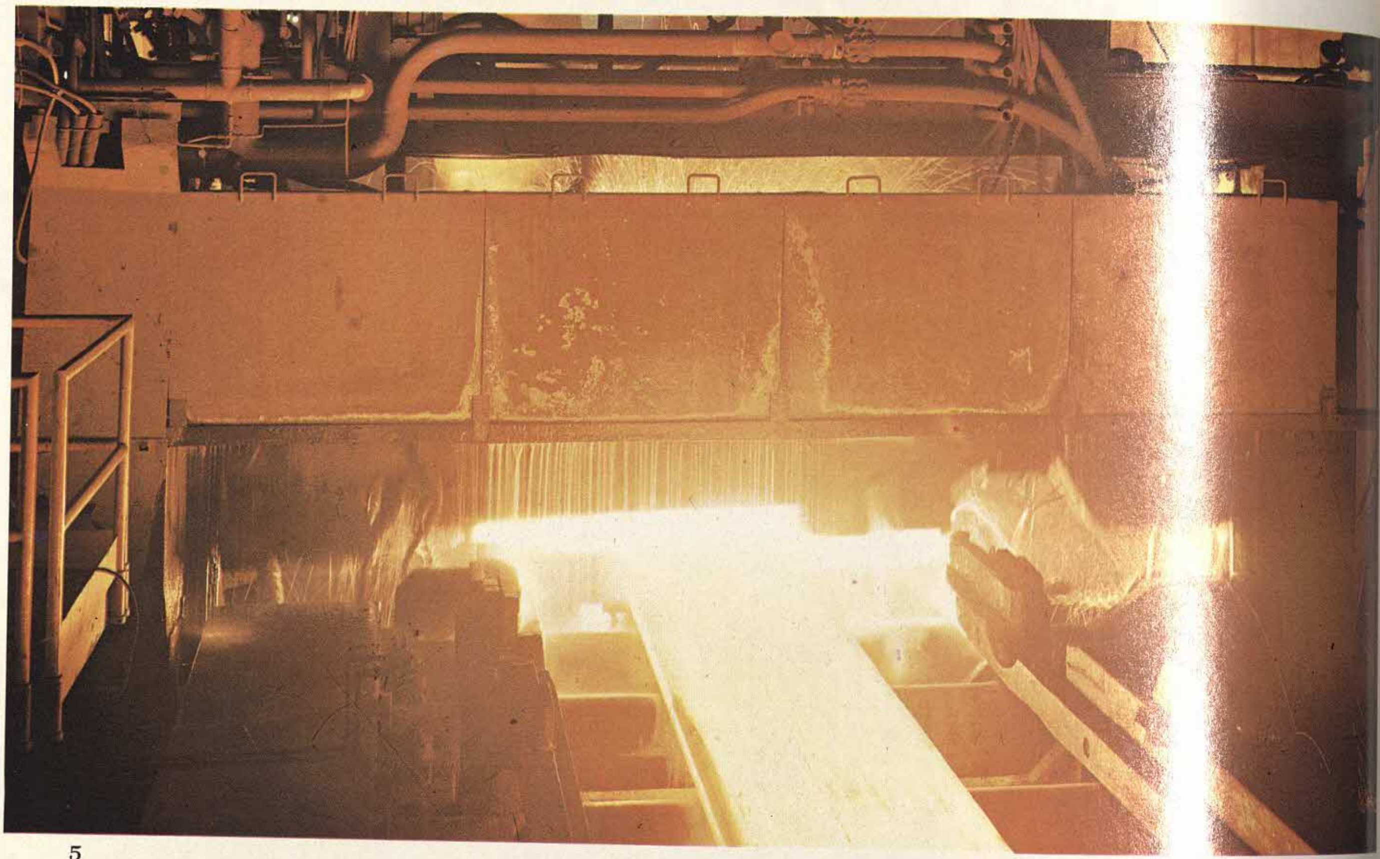
c

powering the train is capable of reversing the direction of rotation of the rollers. The lower of the two rollers is normally fixed, while the upper roller can be raised or lowered vertically by a hydraulic system. In this way, and with the aid of special grooves in the rollers themselves, the ingot can be given the desired thickness and form. Illustration 3c shows an ingot already assuming the form of a flat bloom. In Illustration 3d, the thickness of the ingot is being reduced as it passes between the rollers from right to left.

d

ELIMINATING THE CRACKS—Cracks are the principal defects of steel ingots. Unless remedial measures are taken, small cracks or fissures on the surface of the ingot are transferred almost unchanged to the flat bloom. They can be eliminated by the special machine shown in this photograph. A strong jet of gas and oxygen directed onto the bloom welds the cracks perfectly.

4



5

CUTTING THE FLAT BLOOMS—Immediately after passing through the crack-eliminating machine, the flat bloom enters a large shearing machine, which trims it to exactly the required dimensions. The photograph shows the shearing machine cutting off the rough end of a bloom. The waste pieces drop into a pit below, where they are cooled by a jet of water. They are then removed to the scrap-steel storage, where they are added to other scrap steel to form a new charge for the open hearth furnace or the Bessemer converter.

6

FLAT BLOOM STORAGE. The processing of the ingots is interrupted [illegible] this point. The flat blooms are pushed onto [illegible] conveyor (Illustration 6a), which carries them toward the flat bloom storage. At the storage, they are picked up by a specially equipped crane and stacked into orderly heaps (Illustration 6b). The heaps are then sprayed with strong jets of water to eliminate as rapidly as possible the strong heat they radiate. Later, these blooms will be sent to reheating furnaces to be prepared for passage through the finishing rolling mill.

a

b

THE STROBOSCOPE | making an instrument for observing rapid movement

A falling body, a flying insect, or a drop of liquid falling through the air moves too quickly for the human eye to detect its speed, the precise direction of its movement, or its exact shape. If a revolving object (such as a top, a wheel, a gear, or a propeller) rotates above a certain speed, the eye cannot estimate the speed of movement; it detects only a blur. An instrument that, in effect, can "freeze" (stop) the motion of a moving object can be quite useful. Such an instrument is the stroboscope.

HOW TO OBSERVE A RAPID MOVEMENT

Suppose that a table tennis ball is suddenly dropped from the hand. The unaided eye can see it fall slowly at first and then more quickly, but the eye cannot follow the development of this phenomenon clearly. One way of making such an observation is to drop the ball in the dark, illuminating it briefly at regular intervals. The ball is then visible at different points separated by increasingly greater distances because, as it falls, its velocity increases.

An intermittent light source is even more useful for observing circular motion. If, for example, a six-spoked wheel turns once every 0.1 second, and it is illuminated every 0.017 second, it appears to be almost stationary. The light illuminates the wheel at some given instant and the wheel is observed in a certain position. The light goes off and comes on again 0.017 second later. In the meantime, the wheel (making ten turns a second) moves through one-sixth of a circle. Each spoke takes the position occupied by another spoke during the previous observation. Because all the spokes are identical, an observer is unable to determine whether or not the wheel has moved; it appears to be stationary when it is actually moving quite [illegible].

The same result can be obtained with different illumination frequencies (illumination 0.1 second, for example), but then the image of the wheel flickers. In fact, the light must come on at least 16 to 18 times a second to produce a fixed image. If the speed of the wheel is unknown, the frequency with which the light is switched on can be varied slowly until

THE PARTS ARRANGEMENT OF THE INSTRUMENT—The light source **a** consists of a light bulb fixed in such a way that its filament appears as a straight line.

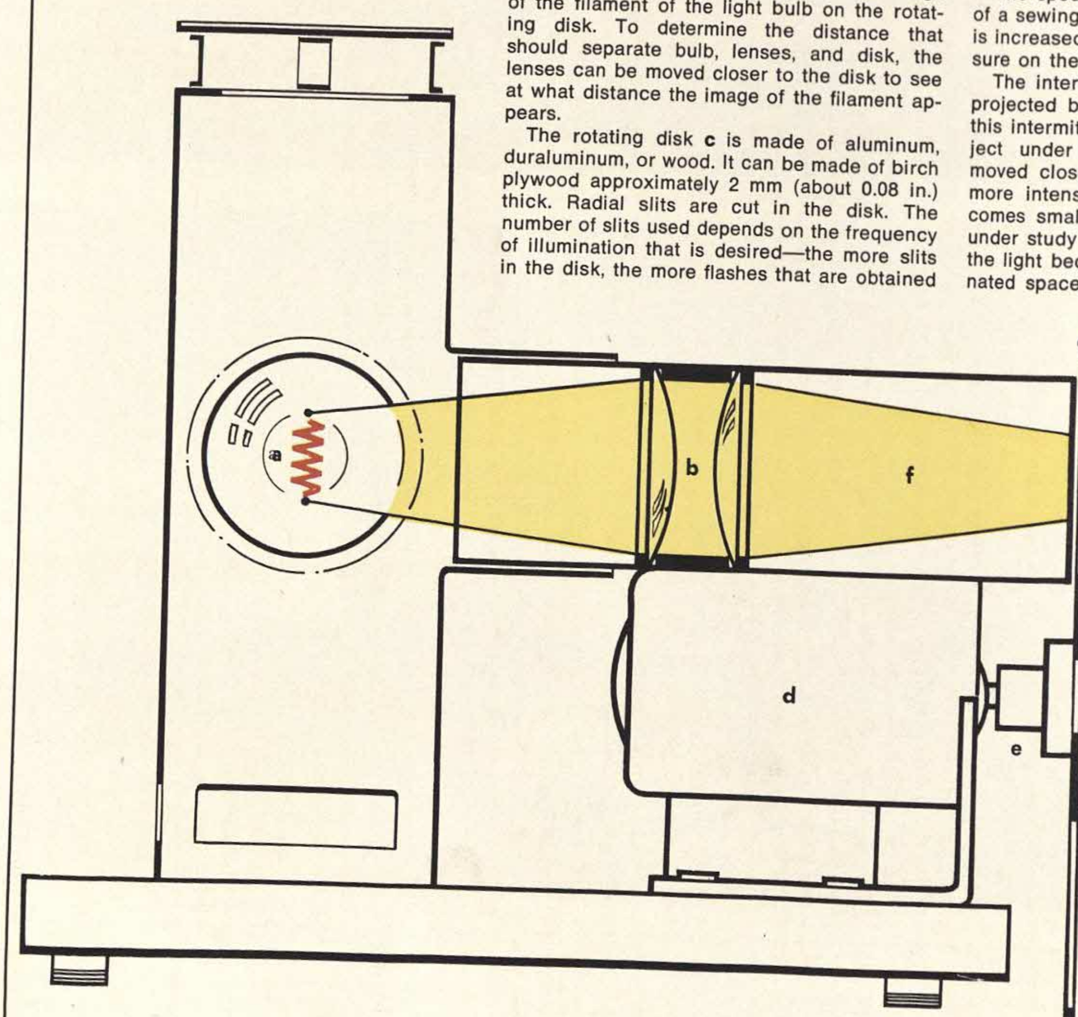
The lenses **b** consist of a pair of plane-convex (or meniscus) lenses with a diameter of 40 to 60 mm (about 1.6 to 2.3 in.) and a focal length of about 60 to 120 mm (about 2.3 to 4.7 in.). The lenses are attached, one in front of the other, almost in contact and with their convex sides facing inward. The light bulb must be mounted at the focal point of the nearer lens. The two lenses thus form an image of the filament of the light bulb on the rotating disk. To determine the distance that should separate bulb, lenses, and disk, the lenses can be moved closer to the disk to see at what distance the image of the filament appears.

The rotating disk **c** is made of aluminum, duraluminum, or wood. It can be made of birch plywood approximately 2 mm (about 0.08 in.) thick. Radial slits are cut in the disk. The number of slits used depends on the frequency of illumination that is desired—the more slits in the disk, the more flashes that are obtained per unit of time and the higher the frequency of appearance.

The variable speed motor is labeled **d**. An old sewing-machine motor is very effective, because it uses little current and easily reaches about 3,000 rpm (50 [illegible]

The speed control device **e** usually consists of a sewing-machine pedal control. The speed is increased or decreased by varying the pressure on the pedal.

The intermittent light is the cone of light projected beyond the stroboscopic disk. It is this intermittent light that illuminates the subject under observation. As the subject is moved closer to the disk, the light becomes more intense and the illuminated space becomes smaller; conversely, when the subject under study is moved away from the opening, the light becomes less intense, but the illuminated space becomes larger.



the spokes of the wheel appear in the same position. Only then does the wheel appear to be stationary. This method of observing periodic phenomena—phenomena (such as the motion of a spoked wheel) that repeat at constant intervals of time—is called the stroboscopic method of observation. The stroboscope, therefore, consists basically of a light source capable of producing an intense light at periodic time intervals.

THE CHARACTERISTICS OF THE STROBOSCOPE

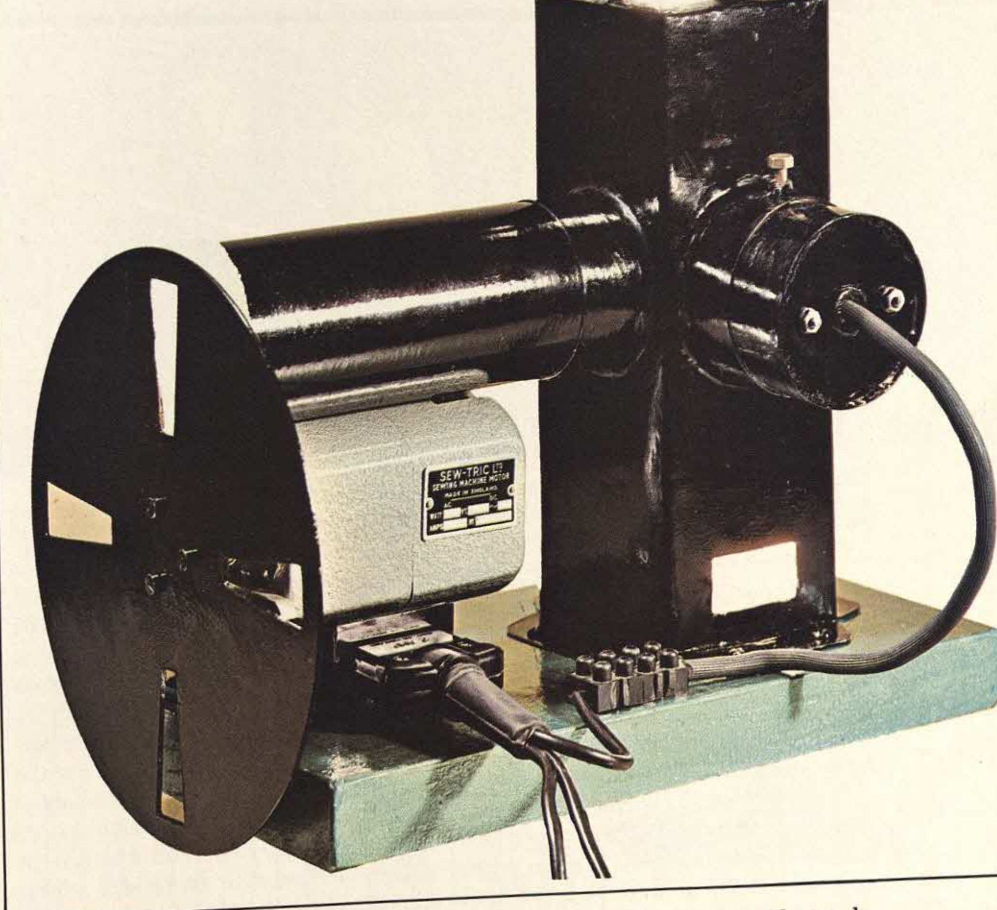
A stroboscope must emit an intense light for a very brief time, and it must emit the light at constant intervals; that is, intervals that occur after equal periods of time have elapsed. It must also have the capability of varying the frequency at which the light is emitted. An example of the construction of an extremely simple stroboscope can contribute to an understanding of the characteristics of the instrument. A stroboscope can be constructed from an ordinary light bulb (for the light source), a revolving disk with slits (to interrupt the rays of light intermittently), and a variable-speed electric motor (to generate the variable light frequencies).

SOME DETAILS OF CONSTRUCTION

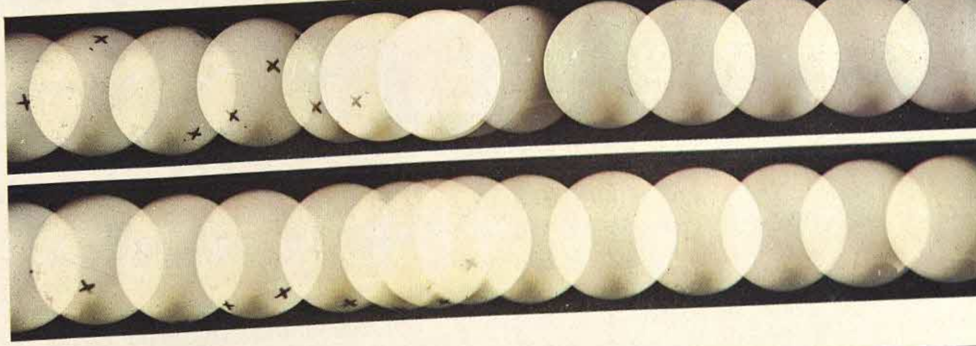
The lamp. The light bulb must have a filament, and, ideally, the filament must be arranged as nearly as possible in a straight line. Not all bulbs have straight filaments, but an absolutely straight filament is not essential, so an arc-shaped filament is acceptable. The bulb must be positioned so that the filament, when seen from the direction of the slits in the rotating disk, looks straight. This arrangement permits the narrow slit in the disk to let the light through only for the short period in which it passes the axis of the light source. Thus, brief and intense flashes of light can be obtained. The socket of the bulb and the support for the two lenses must be fixed so that their positions do not vary. Normally, the most stable position for them is the base on which the motor is fixed.

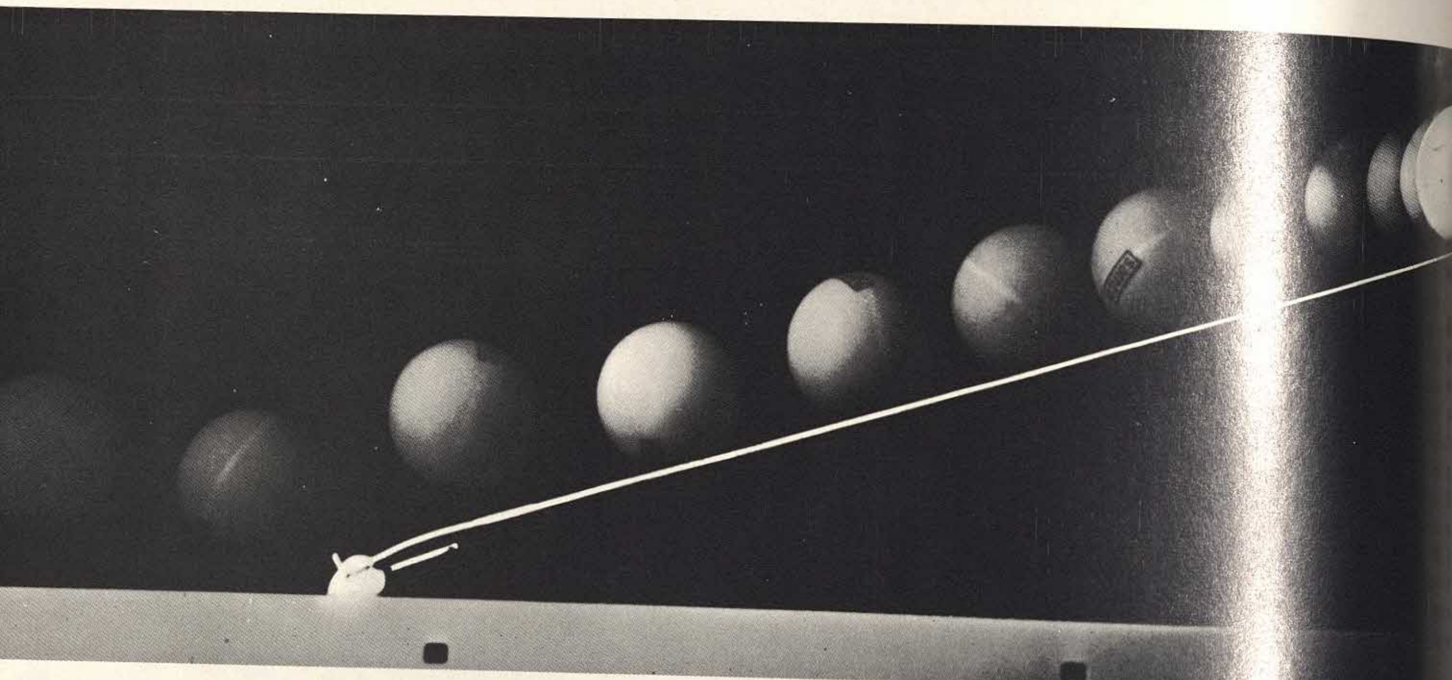
Only the light that falls on the lenses is used. The remaining light dispersed inside the device must be absorbed, and a shield should be constructed for this purpose. A satisfactory shield can be made from a cardboard box with holes in

STROBOSCOPE—A disk containing four slits is attached to the shaft of a small electric motor. Light can pass through the slit opposite the light source. Thus, with each turn of the disk, the object placed for examination in front of the slit is illuminated four times.



PHOTOGRAPHS OBTAINED WITH THE HELP OF A STROBOSCOPE—Two billiard balls are propelled toward each other. By using the stroboscope, an observer can see the various phases of the resulting movement. In fact, each time the objects are illuminated, their new positions can be detected by the human eye or the camera.





MOVEMENT OF A BALL DOWN AN INCLINED PLANE—This type of movement involves uniform acceleration that can be seen clearly with a stroboscope. The illustration shows that, with increases in time, the distance covered by the ball in equal periods of time increases. The increase in speed is constant in time, and the smaller the inclination of the plane, the less increase there is in distance.

it to let the air circulate, cooling both bulb and box. A 100-watt bulb provides adequate light. Bulbs with higher wattage are useful for observation, but require more efficient air circulation and cooling systems to prevent the shield from burning.

The condenser. The lenses of the condenser can be mounted in a paper tube using a cold-drying glue. A glued paper ring can be used to keep the lenses fixed at the correct distance. Wooden wedges can keep the lenses pressed against the ring.

The divided disk. Satisfactory results can be obtained in many situations by constructing only three disks with various divisions. One disk, useful for observing slow phenomena, should have a single radial slit. This disk enables the user to produce flashes of light with a maximum interval equal to that at which the disk rotates. If the motor turns at a maximum rate of approximately 3,000 revolutions a minute, 50 interruptions a second are obtained. The second disk should have about 20 radial slits around it. This disk produces light flashes at a frequency of about 1,000 interruptions a second when the motor is operating at maximum speed. If the slits are narrow, the duration of the flash may drop as low as 0.0001 second, which is lower than the rate that can be obtained from electronic flash units. The third disk should have an intermediate number of divisions, with perhaps five radial slits. The diameter of the disk must not be less than 20 cm (about 7.8 in.) so that the light source can be fixed over the motor. The slits can be slightly longer than the image of the filament of the light bulb, or about 3 to 4 cm (about 1 to 1.5 in.). Their width should be no more than a few millimeters, possibly 5 to 10 mm (0.2 to 0.4 in.). A flanged hub can be used to attach the disk to the shaft of the motor. Other standard accessories are usually readily available, and they may be used to attach the disk to the shaft of the motor.

An electric motor with a variable speed control can be used without modification.

HOW TO MAKE OBSERVATIONS

After the stroboscope has been assembled, the disk should be examined to assure that it is securely fastened to the shaft of the motor so that it does not fall off during the observations. Ventilation around the light bulb should be checked to assure that overheating does not develop. Observations can begin with simple movements such as the fall of a light-colored object that shows up well in the dark when it is illuminated. After a little experience, more complex experiments can be undertaken, such as observations of the beating wings of an insect. In making such observations, the subject is held close to the axis of the light source while the speed of the motor is varied until the insect's wings are "stopped." If the two speeds are only slightly different, the flight is seen in slow motion. Stroboscopic experiments can be visually observed, or they may be photographed for permanent record.

The stroboscope has a great variety of applications for precise frequency determinations and for observations of moving parts. Electric-motor operation can be studied, particularly as to brush action and vibration, as in a vacuum cleaner. In industrial maintenance work, the opertion of gears, cams, and other moving parts can be studied with a stroboscope in order to detect misadjustments, misalignments, and wear before damage occurs. The effect of vibration on components intended for use in aircraft also can be observed under simulated service conditions.

TEFLON | a superior polymer

The best-known fluorine polymer is polytetrafluoroethylene (PTFE), more commonly known as Teflon. The chemical, thermal, and electrical properties of PTFE are superior to those of many other plastic materials. Thus the product meets many of the most rigorous requirements of modern industry.

The interest shown in this polymer is fully justified by its properties: chemical inertia (only metals in a fused state, fluorine at extremely high temperatures, and a few halogenated compounds attack it); excellent resistance to heat up to temperatures as high as 260° C (500° F), and for brief periods, to even higher temperatures; resistance to cold to −265° C (−445° F); excellent dielectric properties; nonadhesiveness; self-lubrication; resistance to aging; no absorption of humidity; and nonflammability.

PHYSICAL PROPERTIES OF PTFE

The PTFE molecule consists of fluorine and carbon atoms bound together by covalent bonds. The bond between carbon and fluorine is one of the strongest known (about 110 kcal/mole). This bond strength is responsible for the excellent chemical and thermal stability of the polymer. The structural arrangement of the atoms in the macromolecule contributes to the exceptional combination of properties in Teflon.

As a result of reciprocal repulsion, the fluorine atoms induce a spiral configuration of the parallel macromolecular chains of the crystalline polymer. The high crystallinity and the consequent compact structure of the polymer explain its insolubility in all known solvents and its high index of viscosity when it is brought to a temperature above the melting point. Repulsion between the fluorine layers, on the other hand, explains the low coefficient of friction.

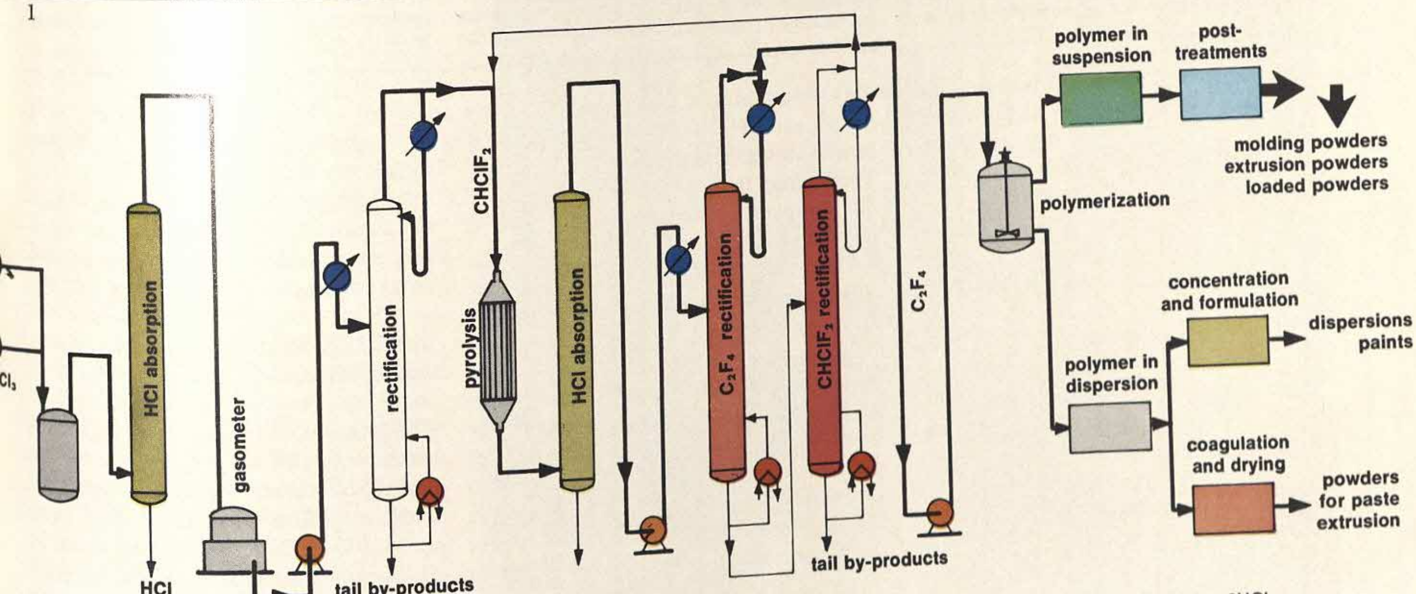
PTFE has various transition points; two of these are important from the point of view of processing. The existence of a transition point at 20° C (68° F) suggests increasing the powder to a temperature slightly in excess of this value, 21 to 25° C (up to 77° F), prior to the start of processing. This ensures that transformation is reproducible as the operating temperature varies to slightly higher or slightly lower values.

The existence of a transition point between the crystalline phase and the vitreous (glasslike) phase at 327° C (about 621° F) requires sintering of the product during processing at temperatures in excess of 327° C. As a result of a low coefficient of thermal conductivity, sintering is caused to occur at temperatures decidedly above the temperature of crystalline fusion. Moreover, this same low coefficient of heat transmission and the accompanying considerable variation of volume require that cooling from the vitreous state take place quite slowly. The long time taken to cross this particular temperature barrier ensures the absence of sharp temperature gradients inside the product. Such gradients could lead to internal stresses or even fractures.

The speed of cooling below the transition point, on the other hand, influences the ratio between the crystalline and the amorphous constituents within the polymer mass. Slow cooling of a sintered piece produces a high degree of crystallinity, while rapid cooling favors the amorphous structure. This is due to the fact that the amorphous matter finds it impossible to recrystallize.

Once the PTFE has been brought to the vitreous state, however, it will never reacquire the high percentage of crys-

1



PLANT FOR THE PRODUCTION OF POLYTETRAFLUOROETHYLENE — The tetrafluoroethylene monomer, from which the polymer is obtained, is produced by cracking difluorochloromethane at about 800° C (1,472° F). This substance, in turn, is produced by reaction between chloroform and anhydrous hydrogen fluoride. The reaction takes place under pressure in the presence of a suitable catalyst. The PTFE production equations are as follows:

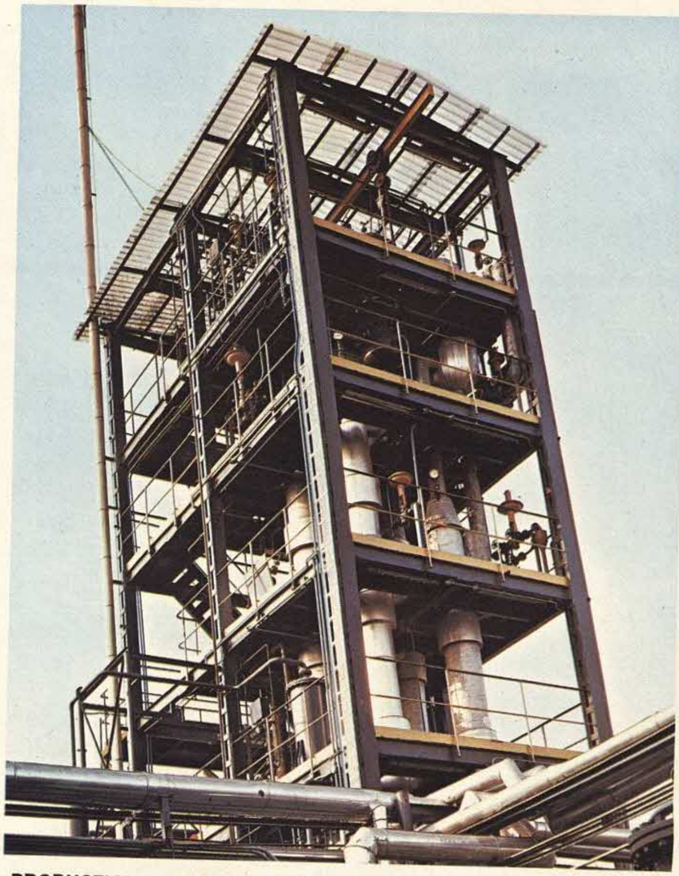
$$CaF_2 + H_2SO_4 \rightarrow 2HF + CaSO_4.$$

$$2HF + CHCl_3 \rightarrow CHClF_2 + 2HCl.$$

$$2CHClF_2 \xrightarrow[\text{or pyrolysis}]{\text{cracking}} F_2C = CF_2 + 2HCl.$$

$$n(F_2C = CF_2) \xrightarrow{\text{initiator}} (CF_2 = CF_2)_n .$$

2a



2b

PRODUCTION OF PTFE — PTFE is produced by polymerization of tetrafluoroethylene. Two fundamental polymerization processes are used to yield two different classes of polymers. In the first—polymerization in aqueous suspension—the resulting polymers have high molecular weights and after appropriate treatments are suitable for molding or extrusion. In the second—polymerization in aqueous dispersion—the aqueous dispersions of these polymers have lower molecular weights. After concentration and formulation, they can be applied as dispersions or paints to various materials by means of soaking, spraying, or brushing. Solidified polymers produced from the dispersions are suitable for paste extrusion. This can be used for producing thin articles, such as bars and tubes with small diameters.

These photographs show the distillation and rectification columns of the monomer (Illustrations 2a and 2b, respectively).

tallinity (greater than 90 percent) that characterizes the polymer at the time of polymerization. This is true no matter what treatment may subsequently be applied to it.

Different ratios between crystalline and amorphous material influence the physical and mechanical properties of the products.

Thermal decomposition of PTFE begins imperceptibly at a temperature of 250° C (482° F). It does not, however, become appreciable for processing purposes until the temperature exceeds 400° C (752° F). Thermal decomposition is then very rapid above 500° C (932° F). The decomposition products are all gases (mostly carbon fluoride, C_2F_4, together with small quantities of perfluorolefins). Therefore, if thermal decomposition were to go to completion, no solid residue would remain.

THE MECHANICAL PROPERTIES OF PTFE

The tensile strength of PTFE is only barely affected by prolonged exposure to high temperatures. Test pieces kept for a month at 250° C (482° F) show a loss of tensile strength of the order of 1 percent. This figure rises to 10 to 20 percent if the temperature is increased to 290° C (554° F).

Deformation of PTFE under load depends on the load itself, the time it is applied, and temperature.

If the stress is high and applied for a very long time, permanent deformation may result. In some cases, however, the product can be restored very nearly to its initial configuration by heating it to above 327° C (about 621° F). Whenever the polymer must be employed under conditions that involve pressures on its surface, however, it is advisable to substitute other materials. The impact resistance of Teflon is high, and will remain rather high even at relatively low temperatures, below −75° C (−103° F). The static and dynamic coefficients of friction are among the lowest to be found in solids, and can be compared only to the coefficients of friction between wet ice and wet ice. These coefficients increase slightly at temperatures below 0° C (32° F), as the load diminishes, and when the relative speed between the contact surfaces is high.

In addition to low values for the coefficients of friction, PTFE is also markedly nonadhesive; in fact, hardly any material adheres to its surface.

ELECTRICAL PROPERTIES OF PTFE

PTFE has exceptional and practically constant electrical properties over the temperature range from −100 to 300° C (−148 to 572° F), and at all frequencies up to 10^8 Hz. Dielectric rigidity increases as thickness decreases. High temperatures have practically no influence on this property even after long exposure.

Volume resistivity remains unaltered even after prolonged immersion in water, and surface resistivity remains unchanged even in 100 percent relative humidity. Teflon's dielectric constant is the lowest known among insulating materials. Its value does not vary for frequencies up to 10^9 Hz and temperatures up to 300° C (572° F) even after long exposure.

The polymer has good resistance to an electric arc flame. Even after prolonged exposure the material decomposes into vapors and leaves no carbon residue. Consequently, its surface resistivity is not modified. Exposure to high temperatures for long periods of time has no effect on surface resistivity. PTFE is, therefore, a material without equal for electrical applications in which its insulating properties are conditioned by high environmental temperatures.

TYPES OF PTFE

Various types of PTFE can be produced by varying certain production parameters and by carrying out specific postproduction treatments. These types are classified according to whether the final product is obtained by molding or extrusion. One or the other, together with a suitable polymer, is used to produce semifinished products at the lowest cost consistent with the properties required for the final applications. Basically, therefore, PTFE materials are divided into molding powders and extrusion powders.

Extrusion powders can be distinguished from molding powders in terms of their greater smoothness; this facilitates automatic feeding of production machines.

In PTFE applications where it is necessary to enhance certain physical properties, such as resistance to deformation under load, rigidity, resistance to wear, and hardness, it is possible to use powders containing inorganic fillers capable of resisting sintering temperatures, which are 370 to 380° C (698 to 716° F).

These charged powders can be molded or extruded. Some require pretreatment, however, to increase smoothness before being extruded. Processing of such charged powders is carried out in a manner similar to the processing of normal powders. Higher working pressures, however, must be used to keep the porosity of the finished products to a minimum.

Powders for paste extrusion are white powders consisting of aggregates with average dimensions of the order of 500 μm (microns). These are obtained by means of coagulation from PTFE dispersions. These powders are extruded in paste form; that is, they are mixed prior to processing with organic lubricants (light naphtha, xylene, solvent gasoline, and so forth). This mixture is preformed in cylindrical dies at pressures of 20 to 35 kg/cm^2 (about 284 to 498 lb/in.^{2}). The preformed material is then placed in a piston extruder and the extruded articles are dried and sintered.

This particular processing technique makes it possible to arrive at extremely low reduction ratios (the ratio between the cross-sectional areas of the extruded article and the preformed material), and to obtain extremely thin finished articles, including round bars, small tubes, various sections, strips, and so forth.

Dispersions and paints are obtained by concentration of a PTFE dispersion, followed by formulation. The principal types follow.

Dispersions for impregnation contain about 60 percent polymer by weight and are stabilized by the addition of a nonionic wetting agent. They find application in the impregnation of porous materials, including glass and asbestos cloths and braids, porcelain, carbon, and graphite. Dispersions and paints for surface linings are suitably formulated for lining aluminum pots. They can be applied by brush, after dilution with distilled water, or by spraying after dilution with distilled water. This is the familiar Teflon-coated cookware.

Paint designed to be used as a primer is an aqueous dispersion containing about 55 percent polymer by weight. Such dispersions can be adapted for application to various metals by the addition of an appropriate quantity of a particular acid solution. Primers produced this way bond well to the surface to be treated. Other paints, consisting of about 43 percent polymer in aqueous dispersion, are formulated to provide a transparent layer on a surface already treated with primer.

A last type of paint, consisting of about 60 percent solid polymer in aqueous solution, serves both as a bonding primer and a finishing paint. It is applied after mixing with an appropriate quantity of the same acid solution used with the primer.

PTFE PROCESSING

Although PTFE is classified as a thermoplastic resin, it differs from other such

THE FINISHED PRODUCT—After the previously described processes and the necessary controls, the polytetrafluoroethylene is ready for processing. The illustration shows the loading of polymer into tank cars.

3

Properties of PTFE	*Value*
absolute specific gravity	2.1-2.3 g/cm^3
specific heat	0.25 cal/g-° C
elongation at failure	180-300 kg/cm^2
hardness	50-60 Shore D
dielectric rigidity	23,000 V/mm
volume resistivity	$>10^{16}$ ohm-cm
surface resistivity	$>10^{15}$ ohm
dielectric constant	2.0
resistance to an electric arc	$>$200 sec

resins because of its very high viscosity in the vitreous state above 327° C (about 621° F) and its complete insolubility in all known solvents.

As already mentioned, PTFE can be processed by applying various techniques —by molding, by extrusion, and by paste extrusion. The molding technique, by means of which it is possible to produce plates, round bars, sleeves, and so forth, is akin to the processing of metal powders or the molding of ceramics. This process consists of preparing a preform by compressing the cold powder at 100 to 300 kg/cm^2 (about 1,422 to 4,267 lb/in.^{2}), sintering the preform (either free or in the mold, with or without pressure) in a furnace with forced hot air at 360 to 380° C (680 to 716° F), and then cooling it at an appropriate rate. The cooling technique affects the characteristics of the finished article, the tolerances, the possibility of producing complex shapes, pieces with inserts, and so forth.

Extrusion is used for producing round bars, tubes, and sections in continuous lengths. Vertical piston extruders are used, as well as horizontal ones with Archimedean screws. In these machines the powder is forced through a suitably heated mold in which the preforming, sintering, and cooling phases are carried out as part of a single continuous process. The extrusion speeds are lower than for normal thermoplastic resins. This is due to the low coefficient of thermal conductivity of the PTFE and the extremely high temperatures that the polymer must reach.

Paste extrusion can be used for producing thin articles, particularly bars of small diameter, thin pipes, tubes with small diameters and thin walls (used for the most part in electrical insulation, nonsintered tapes for gaskets and insulation, and so forth).

As mentioned, this technique uses powders mixed with suitable organic lubricants to facilitate the extrusion. In view of the high specific surface of the powder, the lubricant must be completely adsorbed. The powder is then compressed in a cylindrical mold at 20 to 35 kg/cm^2 (about 284 to 498 lb/in.^{2}) pressure to obtain a preform, which is transferred into the chamber of a piston extruder operating at slow speed. The preform is then extruded through a mold maintained at room temperature, with the sole exception of the exit area, which is slightly heated to about 40° C (104° F). A complete stroke of the piston extrudes an entire preform. A new preform is then welded onto the small residue of the previous one. The extruded product then passes continuously through one or more tubular furnaces in which the lubricant is evaporated and the product is sintered. The output from this type of processing is conditioned by the speed with which the extrusion lubricant can be removed, and also by the rate at which the extruded product can be sintered.

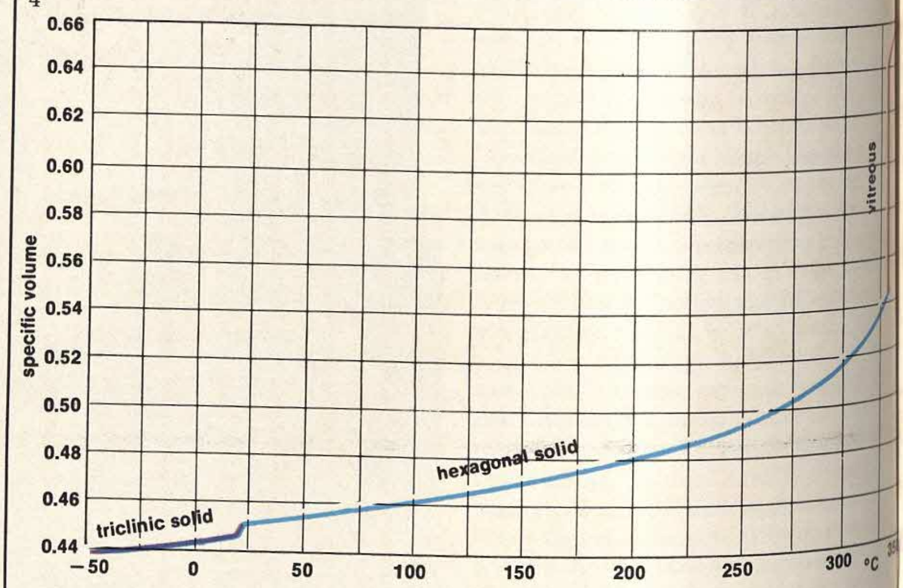
Minimal quantities of gaseous decomposition products containing a toxic fraction may develop at temperatures reached during the sintering process. If these products are inhaled, they can give rise to influenza-type symptoms that will generally disappear within a few days.

SPECIAL TREATMENTS OF PTFE

As already mentioned, PTFE has pronounced nonadhesive properties. These properties are quite useful when applied to hoppers, mixers, reactors, rollers, and so forth. On the other hand, the finished products are quite difficult to glue to other materials. The surface of the polymer has to be specially prepared before gluing becomes possible. In practice this is done by chemically attacking the surface with fused metallic sodium, one of the few chemical agents that corrode PTFE. This leads to a partial defluorination of the surface of the semifinished product. As a result, the surface becomes microporous and readily bonds to glue. The glues used are generally of the epoxy type. The surface of the polymer becomes dark after this treatment because free carbon atoms are formed.

PTFE can be welded by subjecting the carefully cleaned surfaces to the action of temperature and pressure. The surface on which the weld is to be effected must be heated to the sintering temperature and at the same time subjected to pressure. Better results can be obtained by interposing some nonsintered tape obtained from paste extrusion powders. Semifinished PTFE products can be machined to within tolerance limits of ±0.05 mm (about 0.002 in.) with equipment and machine tools normally used for machining metals, wood, and other plastic materials. Special liquids such as water or emulsified oils are not required because the polymer has self-lubricating properties. The use of a refrigerant, however, permits an increase in the cutting speed.

4



THE TRANSITIONS OF PTFE—A change in the crystalline structure of PTFE occurs at 20° C (68° F). The polymer is triclinic below this temperature, but hexagonal above. This change involves a one-percent variation of the specific volume. This is due to the fact that as the temperature increases, a partial loosening of the spiral structure of the macromolecule occurs. The structure passes from a repetitive unit consisting of 13 ($-CF_2-$) monomers to a repetitive unit made up of 15 ($-CF_2-$) monomers. At 327° C (about 621° F) the transition from the crystalline to the vitreous (glasslike) structure occurs. The crystalline structure disappears when this temperature is exceeded. Instead, a translucent vitreous substance accompanied by a 25-percent increase in specific volume is produced. The vitreous substance has an extremely high viscosity, however, and behaves almost like a solid.

A special heat treatment must be applied when small dimensional tolerances are required. This consists of heating the semifinished material to a temperature about 40 to 50 C° above the maximum temperature to which the finished article will be exposed when in actual use and maintaining it at this temperature for a period of time proportional to the thickness. In general, this time period amounts to one hour for every 10 mm of thickness. Cooling must then take place slowly. This treatment ensures the dimensional stability of the finished product because it eliminates internal stresses that may have existed in the semifinished material.

APPLICATIONS OF PTFE

The excellent chemical, mechanical, electrical, and thermal properties of Teflon enable it to be used in a vast range of applications.

Its resistance to all known solvents and practically all the acid and alkaline products used in the chemical industries, as well as the wide range of temperatures over which it can be used, make it a valuable material indeed. Its use has solved many technological problems that would otherwise have remained unsolved.

Moreover, the use of PTFE leads to a considerable reduction in maintenance as a result of longer part lifetime. By way of example, a number of typical applications are gaskets, flat and grooved grommets, and seals in general (a particularly widely used application consists of PTFE tape for sealing screw joints), valve and pump bodies, pump impellers, internally lined pipes and tubes, porous filters, piston rings for compressors, agitators, faucets, hoppers, sheaths, and so forth. The outstanding electrical properties of PTFE permit it to be widely used in the electrical and electronics industries, especially in miniaturization or in applications in which electronic cables are exposed to high temperatures. Extremely thin, flexible coverings for wires and cables are produced. These can be used at temperatures up to 250 to 280° C (482 to 536° F) and in particularly corrosive atmospheres. Extremely thin tapes (with thicknesses down to 0.5 mm) suitable for coatings that adapt themselves perfectly to the shape of the object to be insulated are also produced.

Some typical applications of PTFE in the electric and electronics industries are insulation and covering of wires, cables, motors, relays, transformers, radar sets, junctions, wave guides, and so forth; printed circuits (using PTFE-copper combinations); hermetic closures for condensers and transformers; distance pieces for coaxial cables; bases for valves and cathode-ray tubes; films for condensers, and many others.

PTFE's low coefficient of friction makes it possible to eliminate lubrication in mechanical parts made from the polymer. The deformability of PTFE under load is reduced by loading the polymer with suitable fillers (glass, bronze, and so forth). Some of the more significant applications in the mechanical field are dry bearings, bearing seals, self-lubricating bushings, piston rings for compressors, gaskets for the stuffing boxes of rotating shafts, chutes, PTFE surfaces mounted on rubber supports, conveyor belts, oil rings, washers, gaskets, parts for pneumatic machinery, gear wheels, valve seatings, mechanical folding devices for paper, feed channels for gluing machines, metering pistons, self-lubricating leaves for springs, bridge bearings, and so forth.

When PTFE products are to be used under extreme conditions of temperature, certain important behavior characteristics of the polymer must be taken into consideration. Unlike other plastic materials, PTFE neither melts nor flows at high temperatures. Products made from it retain most of their mechanical characteristics even at 250° C (482° F). They show a certain flexibility even at temperatures close to absolute zero, and do not become fragile when subjected to compression. PTFE has a much higher coefficient of linear expansion than most metals; this must be taken into account when designing a product. A molded piece will shrink about 2 percent when cooled to −150° C (−238° F). On the other hand, it will expand by about 4 percent when heated to 250° C (482° F).

The polymer can be kept indefinitely at very low temperatures without suffering any appreciable deterioration of its properties. It is one of the few polymers that do not lose elasticity when immersed in liquid oxygen. Among the typical PTFE applications at extreme temperatures there are seatings for ball valves, hydraulic regulators, nuclear controls, fuel pipes, flexible tubes reinforced with metal sheaths, missile parts, jet engine gaskets, piston rings for compressors used in the liquefaction of gases, cryogenic mechanical parts, ski linings, pump diaphragms, and so forth.

Because of its nonadhesive properties, PTFE is also used in many applications in the food, textile, cosmetic, paper, and chemical industries. It also finds application in extruders for noodle products, in hoppers, in calenders, and in mixers.

Its stability and physiological inertness contribute to its widespread use in the medical and surgical fields, particularly for instruments, probes, and artificial arteries.

In space applications, quite apart from its use in electronic instrumentation, the polymer is employed in nose cones for missiles and in linings designed to prevent heat losses.

5

SOME INTERESTING USES—PTFE, thanks to its exceptional properties of chemical resistance and nonadhesion, finds many applications. Illustration 5a shows a thin polymer tape that has been colored by means of pigments in dispersion. Such tape is extensively used as gasketing in valves to ensure a perfect seal. Its lack of adhesion enables dismantling of the equipment rapidly and easily.

Illustration 5b shows some manufactured articles; Illustration 5c shows molds for the extrusion of spaghetti noodles. All of these objects are made from PTFE.

TELEGRAPHY | from Morse code to facsimile transmission

The idea of long-distance communication by means of established signals is probably as old as history itself, but it was not until the eighteenth century that significant service networks were developed, based on visual signal transmission. Large objects in the shape of letters or symbols were hoisted atop towers, where they could be seen from afar through telescopes. This was the first regular telegraph service.

During the nineteenth century, when applications of electric current were first being discovered, the transmission of signals through a conductor by means of electrical impulses was explored. The first such system was conceived by the American inventor Samuel Morse. It is based on the attraction exerted by an electromagnet on an armature, moving it in such a manner that signals can be marked in ink on a continuously moving strip of paper. Morse also devised a code in which dots and dashes correspond to letters and symbols. With this system, messages of any kind can be transmitted in any language. The Morse code was widely used for many years.

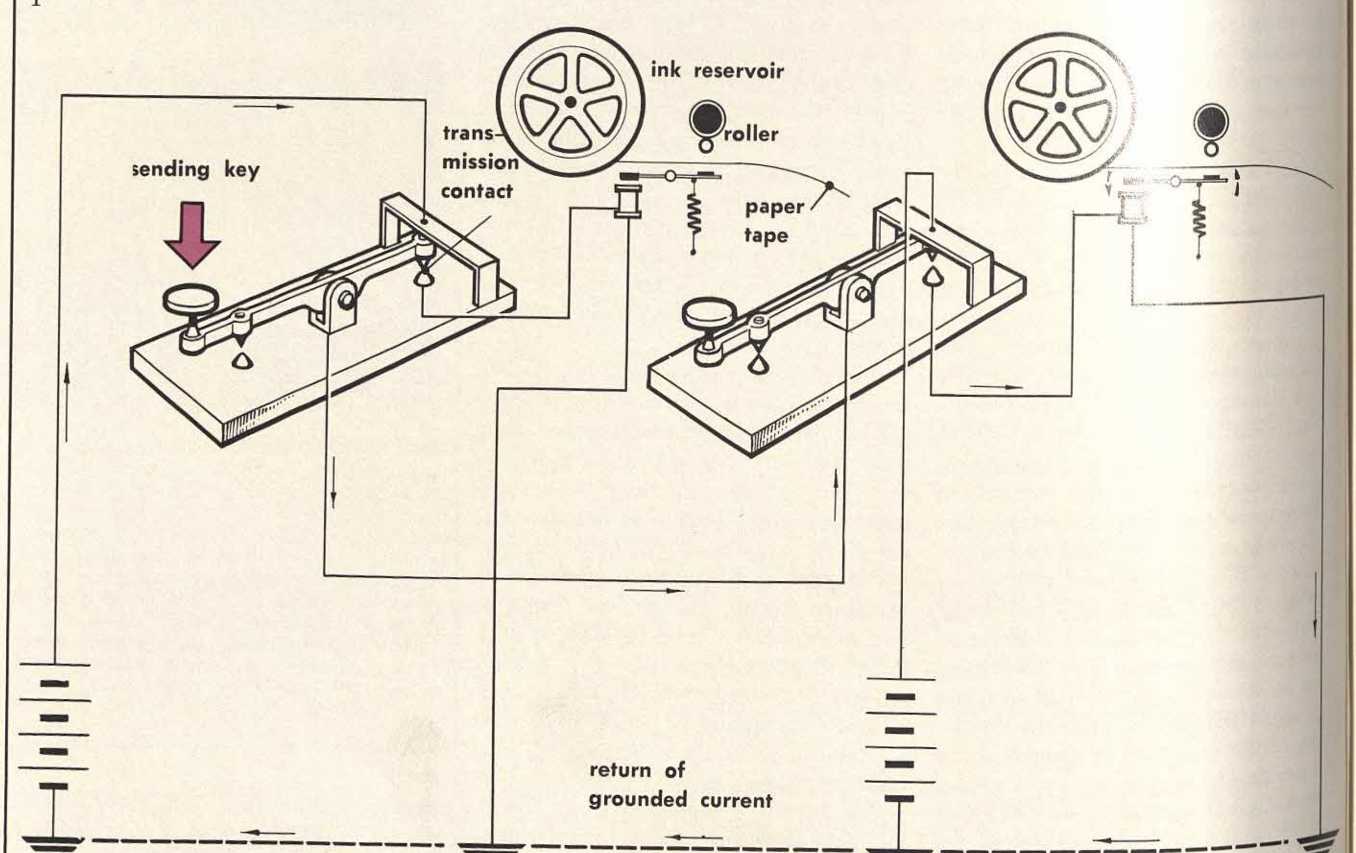
More sophisticated systems were developed later in the century. These include the printing telegraph, which has an alphanumeric keyboard that automatically generates signals corresponding to the particular key that is struck. A similar mechanism at the other end of the circuit receives the incoming signals and prints them. The transmitter is equipped with a continuous impulse emitter that is perfectly synchronized with the receiver, and the sending of signals must follow the rhythm imposed by the synchronizing mechanism. For this reason, the system is classified as rhythmic telegraph.

Along with refinements in telegraph mechanisms came improvements in conductor systems. Prior to this time, the telegraph line consisted of a single wire, with current being returned through the ground. The development of complex synchronization systems permitted the transmission of a number of messages

MORSE'S TELEGRAPHY—Every telegraph station is essentially made up of two parts: the sending key, which transmits the signal, and the receiver. The sending key is movable and has an electric contact point at one end, which alternately closes one of two circuits. The movable part of the key is connected to the telegraph line in both transmitting and receiving stations. When not in use, the closed contact is connected to the receiver's electromagnet and from this point to the ground.

During transmission, the key is pressed, closing a second contact that goes to the battery and has its other pole grounded. The transmitted current lasts only as long as the key is depressed. On the receiving end, the electrical impulse activates an electromagnet that in turn attracts an armature. To allow the recording of signals in the absence of an operator, the armature is attached to a transverse magnetic shell located beneath a paper tape that moves at a regulated speed between two reels. Above the paper are a small roller and an ink reservoir. Each time the armature rises, the magnetic shell presses the moving tape against the roller, leaving a mark on the paper that is proportional to the length of the electrical impulse.

1



by utilizing the intervals caused by rhythmic spacing between one signal and the next.

The introduction of radiotelegraphy considerably enhanced the importance of the telegraph, and greatly extended the use of Morse code, particularly for nautical applications. More recent is the development of arrhythmic telegraphy, which eliminates the necessity of maintaining a pre-established signal transmission speed. This system utilizes teletypewriters, machines that are similar to normal typewriters in that they are equipped with keyboards, typebars, and paper rollers.

A teletypewriter can be used for both transmission and reception. When a key is struck, a "start" signal is transmitted, activating the receiving apparatus; the signal is followed by a coded signal corresponding to the letter transmitted. A "stop" signal shuts off the receiver. This system does not require perfect, continuous synchronization between the transmitter and the receiver, because the "start" signal indicates the precise instant at which the transmission cycle begins.

For teletypewriter equipment, a special five-element code group has been developed. It is made up of a series of signals, either continuous or discontinuous, that can be distinguished from one another because the entire transmission cycle has a predetermined time length beginning from the sending of the "start" signal.

These easily used teletypewriters enjoy widespread use, not only in public telegraph services, but also in many businesses. Commercial users are tied into a vast automatic network, which is equipped with complex switching centers, much like telephone exchanges.

The development of the facsimile technique for the electrical transmission of fixed images, such as photographs and documents, is based on a point-by-point scanning of the subject, a process similar to television transmission but somewhat slower. Because of its slower speed, facsimile transmission operates on a substantially lower-frequency band.

A major use of facsimile is for transmitting a variety of weather maps and oceanographic charts, which are broadcast 24 hours a day. They provide intercontinental weather data to mariners and aircraft pilots.

Satellite facsimile applications have increased from the initial use for transmission of cloud pictures taken by Tiros and Nimbus satellites. Facsimile transmissions of the Ranger, Surveyor, and Mariner satellite missions made close-up views of the moon and Mars.

2

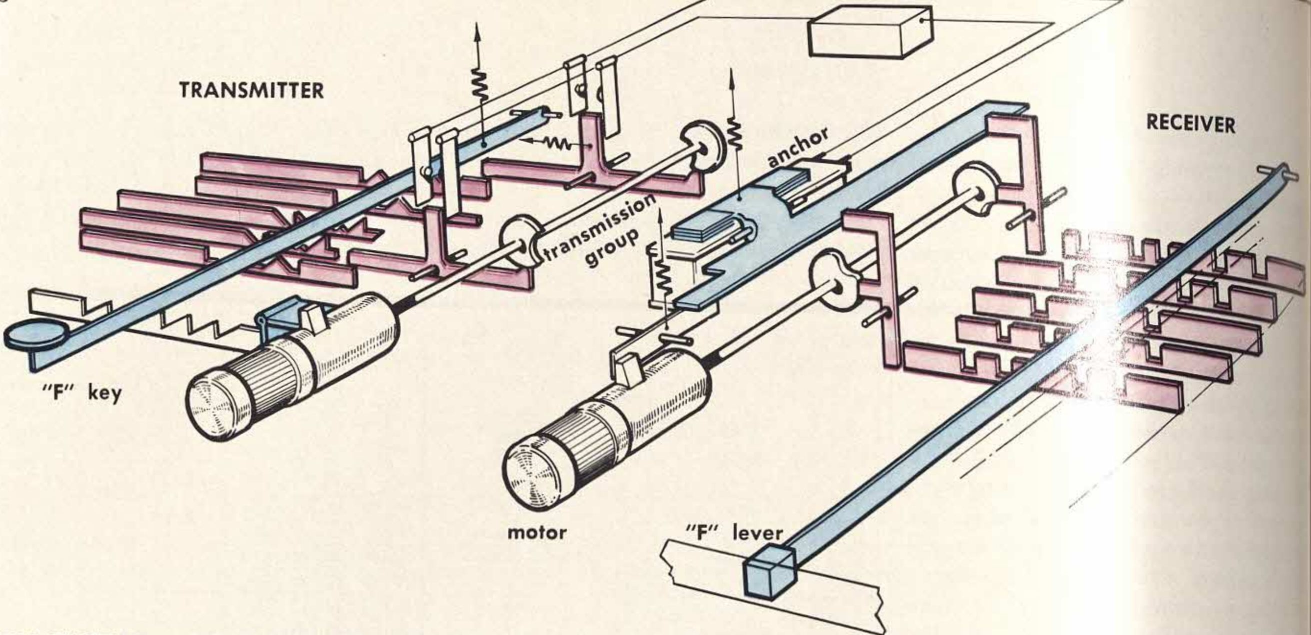
letters	Morse code	numbers, symbols	Morse code	impulse code for telegraphic transmission				
A	.-	—	-....-	●	●	○	○	○
B	-...	?	..--..	●	○	○	●	●
C	-.-.	:	---...	○	●	●	●	○
D	-..	◆		●	○	○	●	○
E	.	3	...--	●	○	○	○	○
F	..-.	◥		●	○	●	●	○
G	--.	◥		○	●	○	●	●
H	....	◥		○	○	●	○	●
I	..	8	---..	○	●	●	○	○
J		⍾		●	●	○	●	○
K	-.-	(	-.--.-	●	●	●	●	○
L	.-..	)	-..-.-	○	●	○	○	●
M	--	.	-.-.-.	○	○	●	●	●
N	-.	,	--..--	○	○	●	●	○
O	---	9	----.	○	○	○	●	●
P	.--.	0	-----	○	●	●	○	●
Q	--.-	1	.----	●	●	●	○	●
R	.-.	4	....-	○	●	○	●	○
S	...	'	.----.	●	○	●	○	○
T	-	5	.....	○	○	○	○	○
U	..-	7	--...	●	●	●	○	○
V	...-	=	-...-	○	●	●	●	○
W	.--	2	..---	●	●	○	○	●
X	-..-	/	-..-.	●	○	●	●	●
Y	-.--	6	-....	●	○	●	○	●
Z	--..			●	○	○	○	●
RC				○	○	○	●	○
AL				○	●	○	○	○
1				●	●	○	●	●
A				●	●	●	●	●
Sp				○	○	●	○	○
◥				○	○	○	○	○
å/á	.--.-	ñ	--.--					
ch	----	ä	.-.-					
à	..-..	ö	---.					
ü	..--	;	-.-.-.					
error	....							
okay	...-.							
end	.-.-							

- ○ impulse pause open contact
- ● impulse current closed contact
- A.... letters switch
- 1.... numbers switch
- Sp space
- ⍾ bell
- RC = < carriage return
- = ≡ line advance
- ◆ = ✚ Who is there?
- ◥ left free for internal use of each country; not usable in international transmission

THE MORSE CODE—Samuel Morse envisioned an alphabet suitable for transmission over his telegraph in which each letter or symbol corresponds to a group of dots and dashes, created by short and long electrical impulses. A brief interval in transmission indicates the end of a letter; a somewhat longer interval indicates the end of a word or a message.

With the introduction of arrhythmic telegraphy and teletypewriters, a different telegraphic code was devised. Each signal is composed of five elements, which can be either current impulses (represented in the illustration by solid black circles) or pause impulses (white circles). As indicated, each combination of symbols corresponds to two signs that can be transmitted: a letter and either a number, symbol, or punctuation mark. To differentiate, the system utilizes a letters switch and a numbers switch. When one of these is depressed, it signals that all transmission following is made up of letters or numbers, until the switching key is again depressed. Other keys in this system are for spacing, line advance, carriage return, and the bell to alert the operator.

3



THE TELETYPEWRITER—The keyboard and typebars of a teletypewriter are similar to those on an ordinary typewriter. In addition, the teletypewriter has the mechanism necessary for the transmission and reception of electric signals.

A teletypewriter is powered by a small motor. When this is switched on, no particular signal is sent over the line, even though a connection may already have been established. Actual transmission begins when a key is depressed, causing a clutch to connect the motor to various moving parts of the machine. These parts go through a complete cycle and return to the rest position. It is during this cycle that transmission takes place.

When a key is depressed, it strikes five traverse bars. Certain of these bars have V-shaped notches; when struck by the key, they are forced to move sideways. Other bars have straight-sided notches, and these remain stationary. The depression of the key also frees a camshaft that makes a complete revolution, allowing each of five electrical contact points corresponding to the five traverse bars to make contact with the telegraph line. Only those bars that have been moved by the key action make contact; the others produce no impulse. Thus, an impulse code corresponding to the letter that has been struck is transmitted.

In the receiver, a similar camshaft is set in motion when the "start" signal arrives. As it turns, it forces the lateral movement of each of the five bars for which a current impulse is received. Thus, the receiving traverse bars are moved or not moved according to the position of the sending traverse bars. Notches in the receiving bars are aligned with the typebar of the letter being transmitted, and the letter is printed on the paper in the receiving teletypewriter.

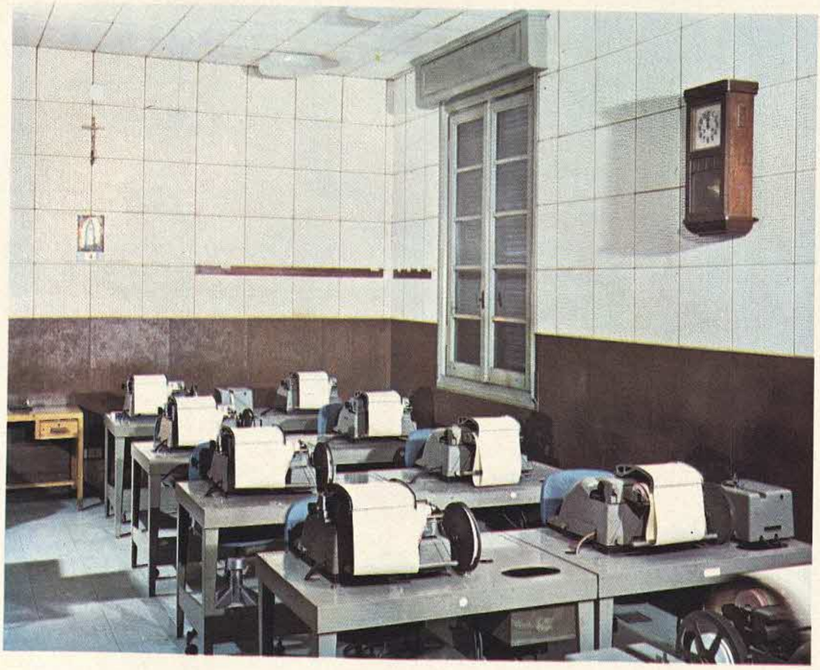
The teletypewriter can also operate with punched tape. Here the code is reproduced by a series of holes. In transmission, the tape permits direct sending of the impulses. On the receiving end, these impulses are recorded by a tape perforator.

An automatic call-sign indicator is often used with teletypewriters. When a given receiving station is called, it automatically verifies that a connection has been made, sending a message to the caller that contains its call sign. This assures that the transmission has been directed properly, even in the absence of an operator at the receiving end.

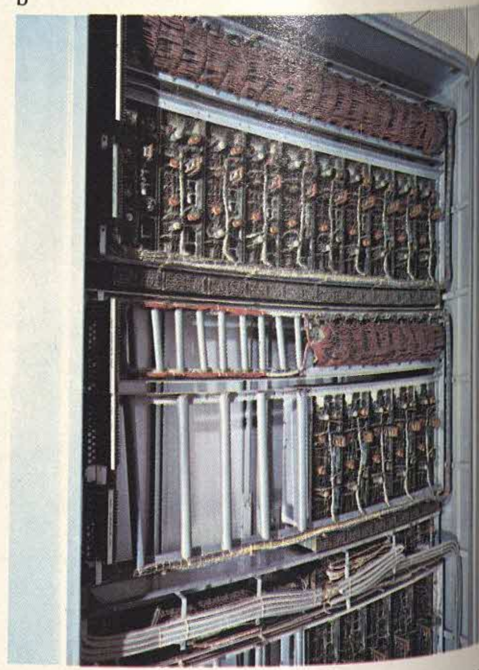
4

THE TELEX NETWORK—Users of one type of teletypewriter service, like public telegraph office customers, can contact one another by means of automatic switching centers, similar to telephone exchanges. The caller simply strikes the number of another user on the keyboard and the connection is automatically established. This system is widely used for international communications. Illustration 4a shows a typical teletypewriter room; Illustration 4b is a detail of the circuit connection panel.

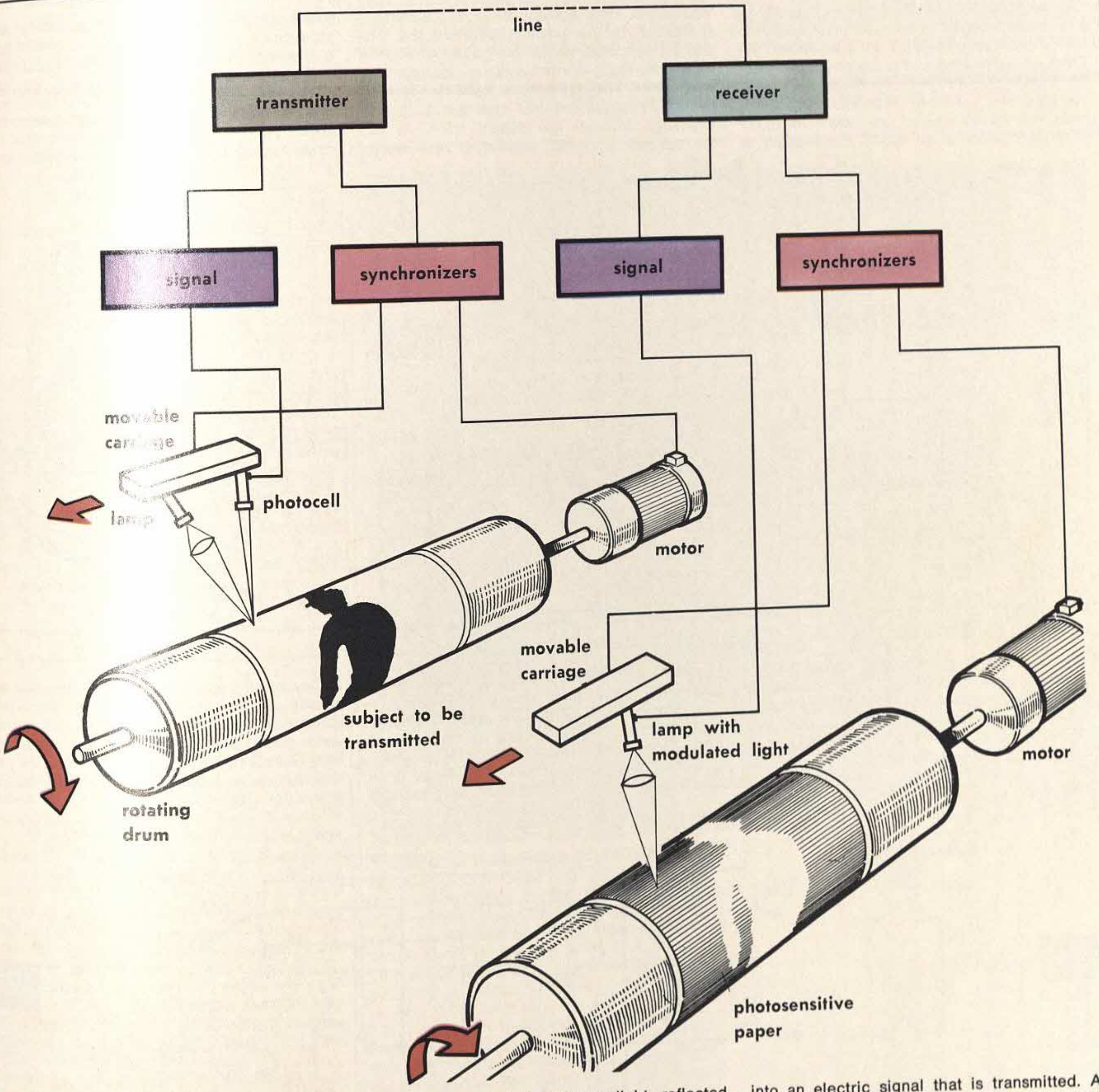
a



b



5



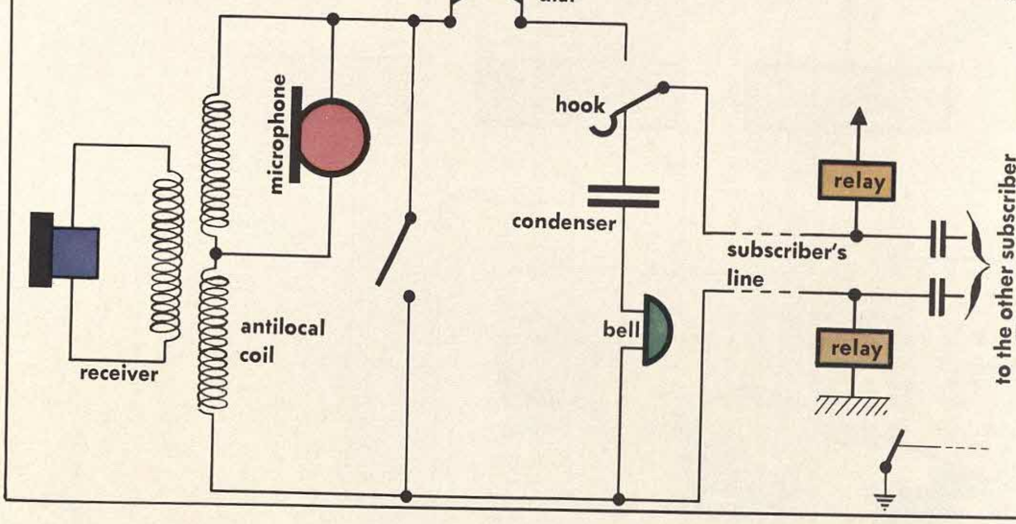
FACSIMILE TRANSMISSION—The technique—utilized also by television—of breaking down an image into points is used for the transmission of photographs, designs, and fixed images over a telegraph line. The item to be transmitted is wound onto a drum that is rotating rapidly at a controlled speed. Opposite this drum is a carriage holding a light source and a photocell that picks up light, reflected from the subject, in proportion to the black-white intensity of the particular point being scanned. The carriage travels from end to end, parallel to the axis of the rotating drum. The two movements provide a total scanning of the subject.

The photocell transforms the reflected light into an electric signal that is transmitted. At the receiver, this signal modulates the light of a lamp that is similarly mounted on a carriage moving parallel to a rotating drum. This drum holds photosensitive paper that is exposed by the light from the lamp. After development, the transmitted image appears on the paper.

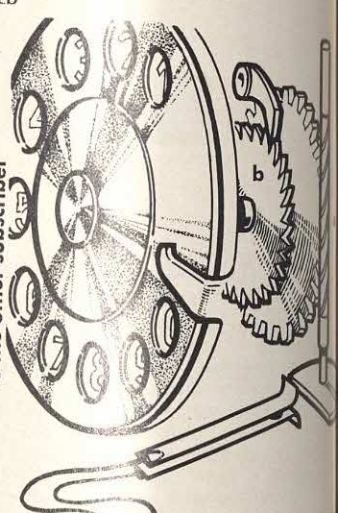
THE TELEPHONE | automatic dialing and electronic switching

THE AUTOMATIC TELEPHONE — The microphone and receiver of an automatic telephone are shown schematically in Illustration 1a. They are connected to the line through a special transformer with three windings. The phonic currents from a distant telephone arrive at the two primary windings and are transferred to the secondary winding, to which the receiver is connected. One terminal of the microphone is connected to the junction between the two primary windings, which are wound in such a way that their magnetic fluxes compensate each other. The microphone itself is operated by the direct current sent from the telephone exchange through the supply relay. In the microphone, a carbon membrane sensitive to sound waves compresses carbon powder and produces a variation in the direct current that is passing. This modulated current is distributed between the two primaries, and a part of it closes locally without affecting the distant telephone exchange. In this way, none of the signals from the microphone reach the receiver, eliminating any local sound effect that

1a



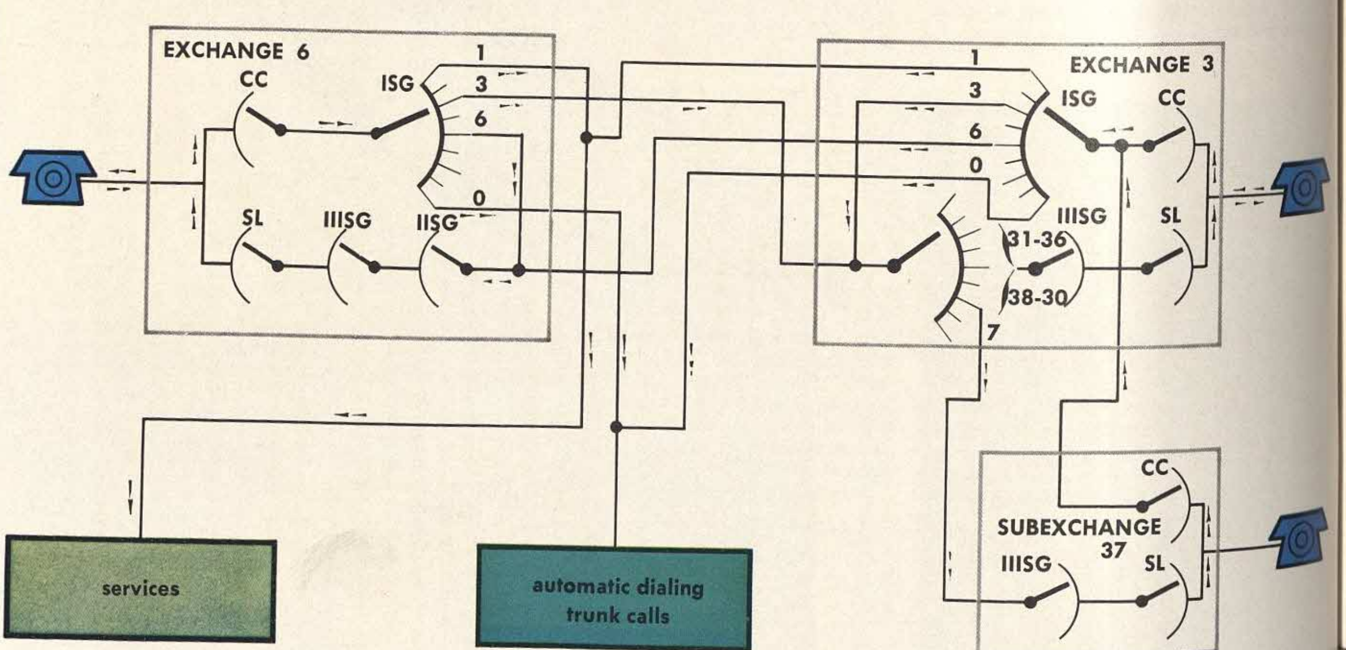
1b



AN AUTOMATIC TELEPHONE NETWORK — The telephones of the subscribers are connected to selectors at the telephone exchange. These selectors are automatic devices that establish connections with other subscribers. Because each telephone must be able to call out or be called, it must terminate at both the incoming and outgoing side, as shown in the illustration.

At **Exchange 6,** on the left, each telephone is connected to the first group of selectors (**ISG,** or Selector Group I). This selector group channels the communication. If the call is to another subscriber within **Exchange 6,** the **ISG** registers the first number (6) and passes the second number to **IISG.** The third number is channeled to **IIISG,** which chooses a line selector (**SL**). The last two digits are received by the last selector (which, in this special application, is commonly called a connector). When the last digit is dialed, the connector

2



would reduce the intelligibility of the communication.

The dialing device (illustration 1b) consists of a rotating cam a—which opens a contact placed in series with the line—and ratchet **b.** The ratchet ensures that the cam cannot move while the dial is being moved forward by hand; during the return motion, however, it is dragged along and opens the line as many times as the number that has been dialed. These interruptions of the continuity of the line cause a corresponding number of actions in the supply relay that transmits the pulses to the selectors at the telephone exchange. In the hook—or receiver rest—are a series of contacts that are commutated by the fork or cradle on which the receiver/microphone unit rests. When the telephone is not in use, the contacts of the rest connect the line to the ringer (through a condenser) in such a way that a 25 Hz alternating current of an incoming call can be received, but no direct current can flow through the circuit. When the receiver is lifted, the ringer is disconnected from the circuit and the microphone inserted. This allows direct current to flow and the supply relay to remain attracted, signaling the exchange that the subscriber is using the telephone. The direct current flows only from the telephone to the supply relay. Immediately after the relay, two condensers permit only phonic currents to pass through the internal lines of the exchange.

rotates to the dialed position.

If a call is made from **Exchange 6** to **Exchange 3** (on the right in the illustration), the first number received by **ISG** is 3. In this case, **ISG** passes the communication to **Exchange 3** and the same series of events described above takes place at the exchange called. If **Exchange 3** is in an area with a large number of subscribers, it is probable that a subexchange (shown in the lower right of the illustration) gathers and processes the calls made to a number of local telephones. In the example shown in the illustration, the subexchange services all subscribers whose numbers begin with 37. Incoming calls with the digit 3 are channeled to **Exchange 3;** when **IISG** at **Exchange 3** receives a number 7, it channels the communication to the subexchange. The subexchange only needs a **IIISG** and a line selector (**SL**), because it processes only the final digits of the number (37—) called.

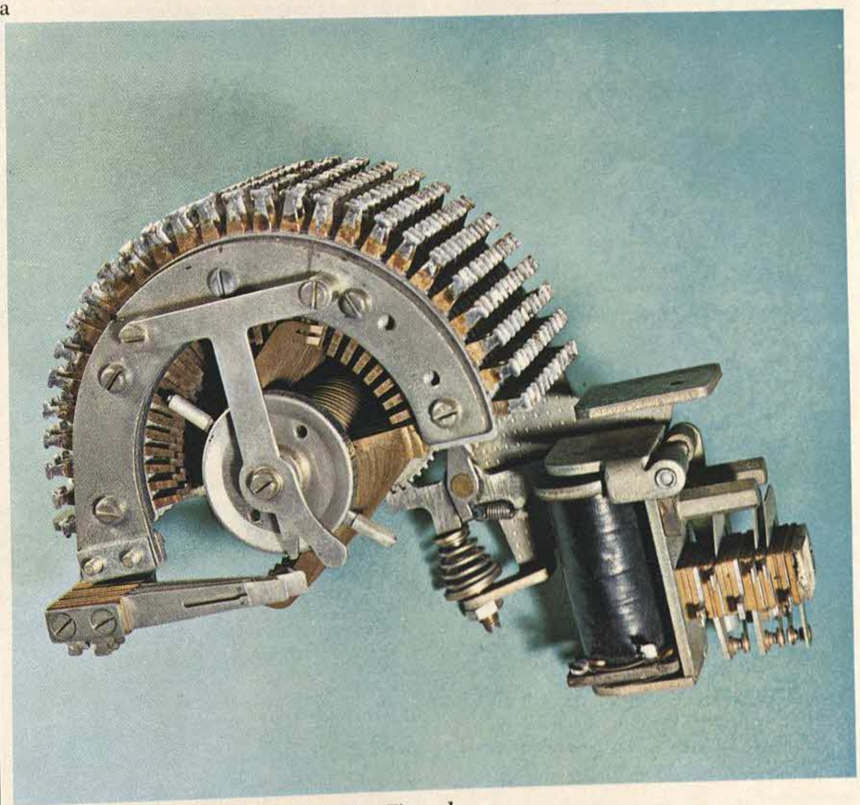
Normally, numbers with the first digit 1 are channeled to the telephone services. Numbers 9 and 0 are used to channel communications to the automatic dialing trunk for short and long distance calls.

There are never as many **ISGs** as there are subscribers, because only a small percentage of subscribers use their telephones at the same time. For this reason, searchers (designated **CC** in the illustration) automatically connect an **ISG** to each subscriber at the moment he wants to use his telephone. On the average, there are 10 to 12 searchers and **ISGs** for every 100 subscribers.

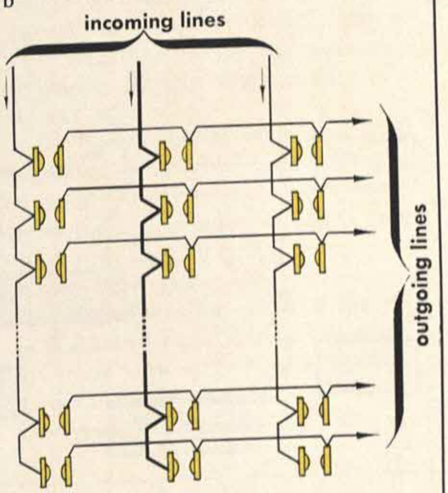
The term *telephone* (from the Greek roots, *tele* "far," and *phoné*, "sound") was first used to describe any apparatus for conveying sounds to a distant point. Specifically, the word was applied as early as 1796 to a megaphone, and not long afterward to a speaking tube. Subsequently the name *string telephone* was given to the device invented long before by the English physicist Robert Hooke,

3

a



b



AUTOMATIC TELEPHONE SELECTORS — The various types of telephone selectors are differentiated mainly by the manner in which they operate. Illustration 3a shows a rotating selector. Rotating (or lifting and rotating) selectors have mechanical arms that rotate (or lift and rotate) to reach a preselected position in a bank of contacts. These contacts are arranged in the form of a circular arc, with 100, 200, or more connections. The movement can occur in various ways but generally is commanded (directly or indirectly) by pulses emitted by the dial. The selector can also search automatically for a free line among ten or more equivalent lines.

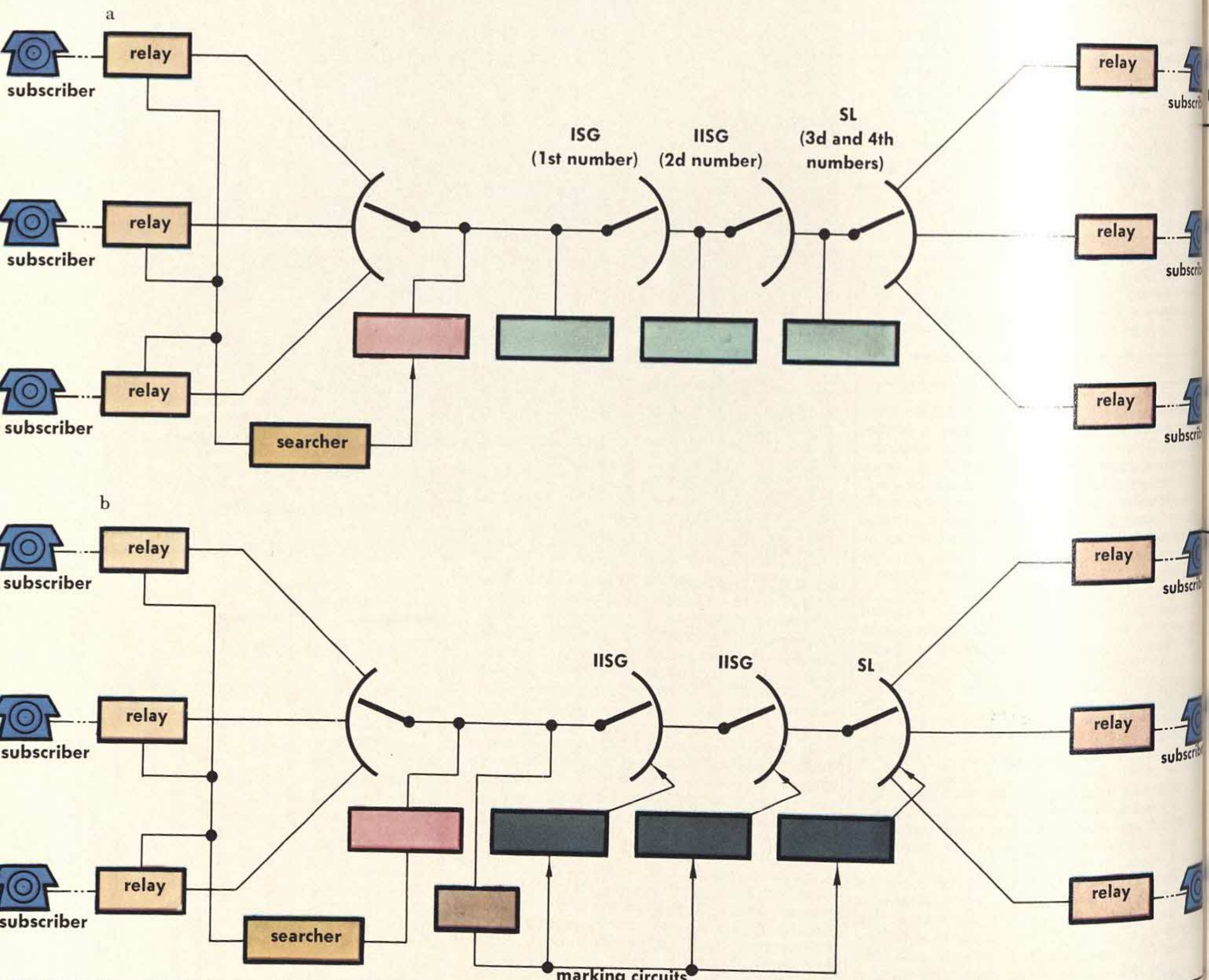
Illustration 3b shows a type of selector called a cross-bar selector. It consists of a matrix of relays arranged in a compact and simple manner. These relays close one at a time, connecting horizontal and vertical lines at their intersections. Each incoming line is connected to a series of contact springs, matching a similar series of contact springs connected to the outgoing lines. Actuation of the relays closes the contacts. This is not done by the dial pulses, but through an intermediate mechanism.

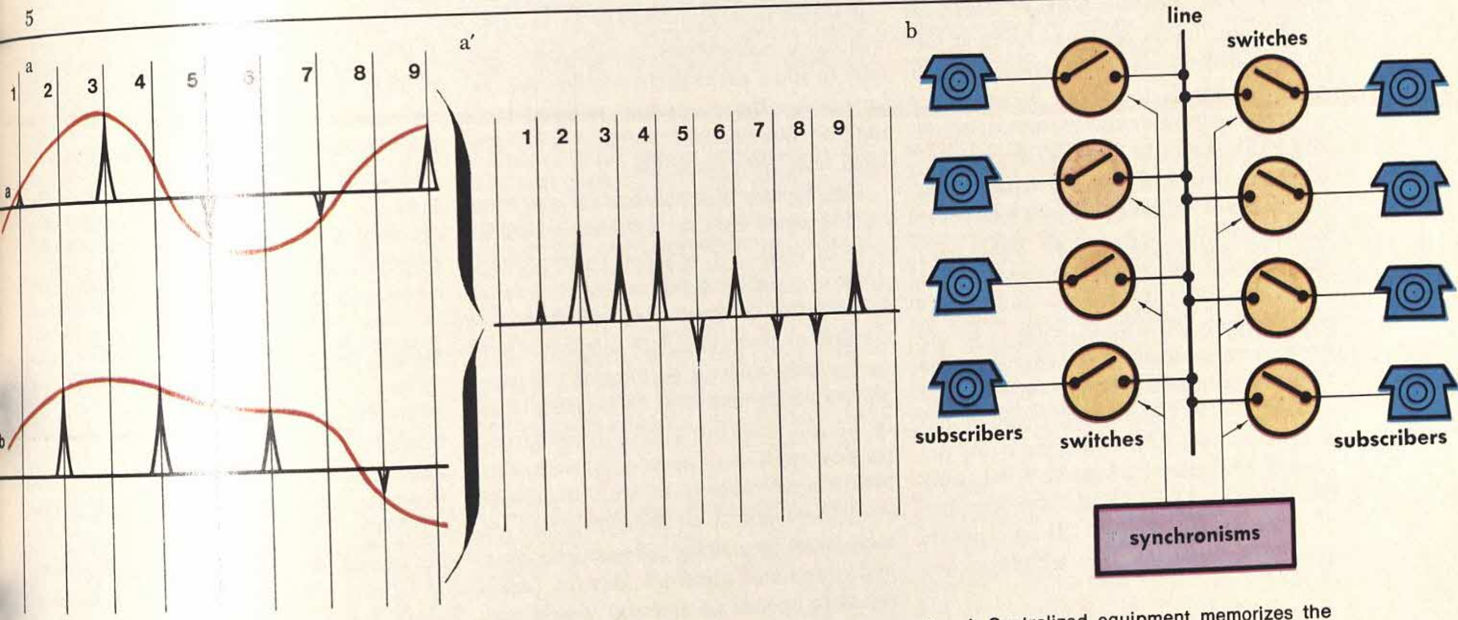
4

TELEPHONE EXCHANGE LAYOUTS — Automatic telephone exchanges are of two basic types: exchanges operated by direct command, and exchanges operated by indirect command.

Illustration 4a shows a direct command exchange. Pulses are received directly from the dials, which control the movement of the selectors, pulse by pulse. Each subscriber's line is connected to two relays at the exchange; these are known as subscribers' relays and show whether his telephone is or is not being used. These relays also serve to set in motion the process of channeling the subscriber's calls through a centralized arrangement that brings the searchers into operation. Each searcher is connected to an **ISG** (see Illustration 2), through which the communication is passed to subsequent selector groups (**IISG**, **IISG**) and thence to the dialed number.

An indirect command exchange is shown in Illustration 4b. This system involves a register that receives pulses coming from the dial, checks them, and identifies the positions to be taken by the selectors (ISG, IISG) to establish the desired connection. The register marks the position of the first selector (**ISG**), which begins to move in selection of this marking. Similar procedures are carried out by the subsequent selectors until the desired subscriber is reached. The method of numbering subscribers in this system is not dependent on the network because the pulses do not directly control the movements of the selectors; instead, they only transmit the instructions from the subscriber to the register.





THE ELECTRONIC TIME-SHARING SYSTEM— This illustration shows a new method of commutation, based on the fact that a phonic current does not necessarily have to be transmitted in its entirety to enable the receiving organs to reconstruct its exact form. A pulse that is proportional to the height of the phonic current curve at that instant can be transmitted at regular intervals of about 125 microseconds. A filter at the receiving end then reconstructs the initial form. If several communications are sent simultaneously, a sample of the various calls—at identical but suitably dephased intervals—permits the transmittal of all the relevant pulses on the same line with no confusion of calls.

For example, Illustration 5a shows two signals, **a** and **b,** transmitted on a single line; Illustration 5a′, immediately to the right, shows how the pulses are transmitted.

Applying this system to practical communications, Illustration 5b shows a time-sharing telephone exchange with a group of subscribers served by the same line. Each telephone is connected to the line through a synchronized electronic switch. The sample pulses of the phonic currents can transmit along the line whenever the switches are closed. Centralized equipment memorizes the data of the two subscribers in communication and assigns them a time channel—a particular instant within the complete cycle of 125 microseconds wherein their sample pulses can be transmitted. Once in each complete cycle, the central equipment closes the two switches for a brief moment, permitting the pulses to pass from one subscriber to another. The two switches remain open for 125 microseconds; other pairs of switches are closed during the same time interval. A filter transforms the pulses of the phonic current into direct current, which is then transmitted to the telephone.

in which vibrations in a diaphragm caused by voice or sound waves are transmitted mechanically along a string or wire to a similar diaphragm that reproduces the sound. Still later, devices employing electric currents to reproduce at a distance the mere pitch of musical sounds were called telephones. Nowadays, however, this name is assigned almost exclusively to apparatus for reproducing articulate speech and other sounds at a distance through the medium of electric waves. The term *telephony* covers the entire art and practice of electrical speech transmission, including the many systems, accessories, and operating methods used for this purpose.

The inventor of the first practical telephone was Alexander Graham Bell. He probably conceived the idea of telephony in 1874 but saw no way to make it work. The following year he was working with his assistant, Thomas A. Watson, on a "harmonic telegraph" that should carry several telegraph messages at once. While conducting an experiment he accidentally discovered how to transmit a musical tone over an electric wire. The device might have remained a curiosity if Bell had not been thinking about transmitting speech by electricity. Within a short time he designed the first electric telephone to carry speech.

The telephone makes use of two processes: transmission and commutation. Transmission concerns the methods and equipment for transforming sounds into electric currents (microphone) and back into sounds again (receiver), and for conveying the currents from one point to another through wire, coaxial cables, or over radio relay.

The capacity of the long-distance telephone network has been expanded by use of coaxial cable and radio relay. Coaxial cable contains as many as eight copper tubes with a wire running through each of them. Plastic disks about an inch apart keep the wire in the center of the tube so that both have a common axis—hence the name coaxial. One pair of coaxial tubes can simultaneously carry some 600 conversations plus two television programs.

The microwave radio-relay system consists of directional antennas on towers about 30 miles apart. A very-high-frequency radio signal is beamed from one to the next. Radio relay can transmit the complicated television frequencies and a large number of telephone messages.

Commutation is concerned with everything that serves to establish or interrupt a connection and for selecting—either manually or automatically—a particular subscriber from among all those connected to a network. The receiver rest (or hook), the bell, and the dialing disk are elements of commutation. The receiver rest establishes or interrupts the connection with the telephone exchange and signals the command to begin or end a conversation. The bell signals the arrival of a communication. The dial sends the signals that identify the number of the station called. These signals are transmitted to the telephone exchange.

EARLY DEVELOPMENT

The first telephone exchanges were manual. Connections had to be made by an operator, at the verbal request of the subscriber. Each subscriber's line terminated at a jack placed on a vertical switchboard. A lamp next to the jack lighted up when the subscriber lifted his receiver. The operator then inserted a plug into the jack to hear the subscriber's request. The requested connection was established by inserting a plug into the jack of another subscriber, this plug being attached to the same cord as the first.

In 1888 the first automatic selector was invented; it was an electromechanical device with small rotating arms capable of exploring a semicircular bank of contacts. This selector was capable of establishing a connection between two lines from a command received by means of electric pulses, without the intervention of an operator. The device led to the construction of the first automatic telephone exchanges, which functioned solely on the basis of commands received from the subscribers through dialing. The functions performed by the telephone exchanges gradually became more complex as the size of telephone networks increased. Included in this development was direct long distance dialing, whereby it is possible to obtain an automatic and direct connection from one city to another and from one country to another. For these connections, a subscriber has only to add a prefix to the number dialed; he obtains the connection without having to communicate verbally with the local telephone exchange.

Developments in the field of electronics suggested that the electromechanical devices might be replaced by electronic devices capable of performing the same functions of commutation, but the introduction of electronics into telephony was particularly difficult because of the complexity of the functions to be performed by an autocommutator. There were two possible solutions. One was to use a combination of devices. In such a system, electromechanical equipment would establish the connection between the lines, and centralized electronic devices (analogous to special computers) would perform all the intellectual tasks of gathering information and evaluating the instantaneous situation of the equipment before any commands were transmitted. The second solution was to use electronic devices exclusively. This solution, however, abandons the classical concept of a spatial connection and replaces it with a concept involving the temporary assignment of connection components. This leads to commutation on a time-sharing basis—typically and exclusively electronic.

THE TRANSISTOR

Foremost in the advancement of the telephone industry is the transistor, a three-electrode amplifying device employing a semiconductor such as germanium or silicon. The transistor can perform many of the functions previously assigned to vacuum tubes, but is many times smaller and much more efficient because it requires no power to heat a cathode. Furthermore, it operates on much lower electrode voltages. Because of the transistor's low power dissipation and low voltage requirements, other electronic components associated with it can be miniaturized. The overall result is that the size, weight, and power consumption of apparatus employing transistors usually can be reduced to a small fraction of that for equivalent vacuum tube apparatus.

Transistors therefore find large and varied applications in almost every field of electronics, and particularly in telephony and associated services, including not only places where vacuum tubes were previously used but in many new situations as well. Thus transistors provide economies through more extensive use of amplification and carrier techniques. (See the articles on transistors in this book.)

HIGH SPEED SWITCHING

In contrast with relays and other electromagnetic devices that require a minimum of several thousandths of a second to operate, electronic devices such as the transistor may be used to perform switching operations with a speed of the order of a few millionths of a second. Besides being faster, the electronic devices are also capable of performing more complex operations involving choices based on conditions existing at the instant. This means that no longer must a large complement of apparatus, or even a large part of it, be set aside to serve a customer during the entire period of his call. The extremely high speed was expected to afford substantial economies in new switching systems through centralizing of functions and time sharing of the apparatus used for establishing connections. The common apparatus would serve a large number of customers in such rapid succession that each one received the equivalent of continuous and exclusive service. Reduction of space and power requirements would yield further economies. Altogether, the new art of solid-state electronics provided the basis for a revolution in both switching and transmission technology.

As to telephone service in general, there are no indications of saturation in demand, even in highly developed areas. Substantial further growth in the number of telephones and their use is therefore probable. Large expansion can also be expected in services associated with the telephone industry—such as those provided by video telephones (opened in limited commercial service in the United States in 1964).

TELEPHONY | words through a wire

The human voice is man's basic means of communication. Modern science and technology have perfected methods that allow man to use his voice for instantaneous person-to-person communication over long distances. High-frequency radio waves traveling through space can be utilized for this purpose. It is as if an imaginary channel were created through which the energy of electromagnetic radiation can flow with little or no dispersion. Signals, telephonic communications, television programs—all can be broadcast along this channel, and the number of signals that can be transmitted increases in proportion to the frequency of the carrier wave. High-frequency cable transmission is similar.

TRANSMISSION SYSTEMS—Multiplex transmission of telephonic communications can take place in any one of three ways, depending on the system used to distinguish between different signals from different channels.

The space-division system (Illustration 1a) is normally used in urban telephone exchanges. Each channel utilizes a single two-strand metal wire that maintains the integrity of the particular channel. This may be attached to a specific telephone or line, or it may be connected by a selector only when a signal is given and a specific destination is indicated.

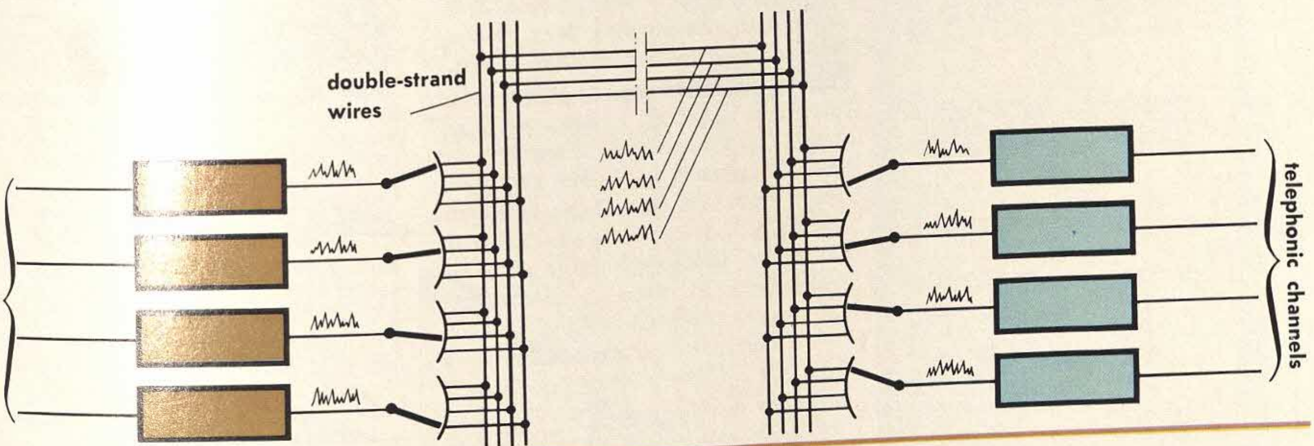
The frequency-division system (Illustration 1b) is normally used in high-frequency transmission. Each channel has its own frequency interval to distinguish it from other channels carried on the same band. Fixed connections are not generally used. Contact is established only when a communication is to be sent.

In the time-sharing system (Illustration 1c), a single channel carries various messages, each kept separate from the others by the variation in transmission time. Sounds create a phonic current of variable intensity. In the space-division and frequency-division systems, these currents are transmitted in their entirety. However, to construct the shape of a radio wave's phonic curve, it is only necessary to transmit the phonic values of this curve at certain predetermined instants that are close enough to one another to indicate all variations in the curve. The variations may be compared with a hospital patient's fever chart. Even though the patient's temperature is noted only once an hour, the progress of the fever can be traced with reasonable accuracy by drawing lines between these hourly notations on the chart.

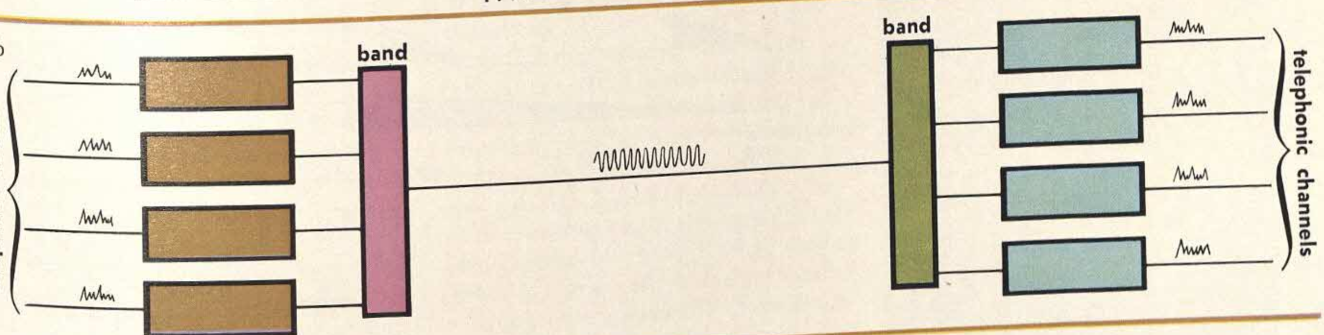
In the time-sharing system, the phonic current is scanned and current values are selected for transmission at regular intervals. Between one value and the next is an open space that is available for use by another phonic current—that is, another channel. The most critical element in this system is the synchronization mechanism, represented in Illustration 1c by a rotating selector, but actually made up of a series of electric currents. This mechanism connects with each of the channels successively, while at the same time maintaining perfect synchronization with a corresponding mechanism at the receiving terminal that makes simultaneous connections with the appropriate channels.

At present, time-sharing systems are extensively used in computer applications. In the future, they will be useful in electronic switching operations.

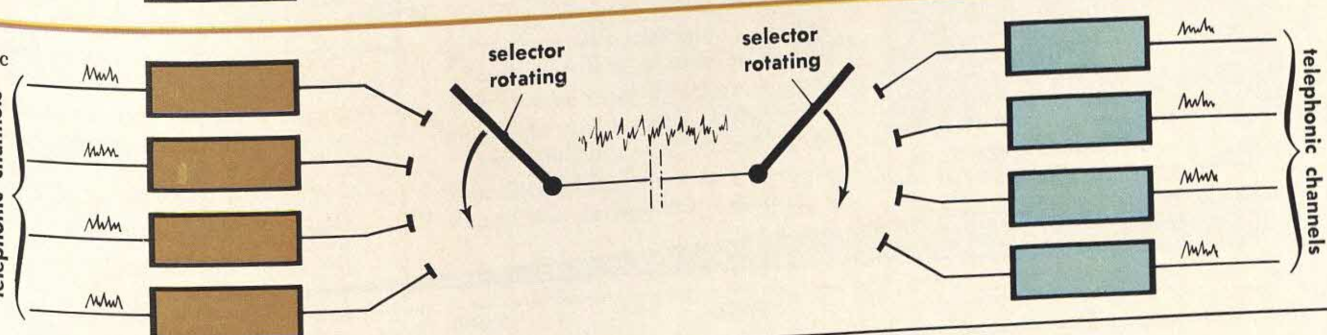
1a



1b



1c



2

FORMATION OF THE TRANSMISSION BAND—In telephony, the phonic band to be transmitted is limited to frequencies between 300 and 3,400 Hz (hertz). While a band of this width (Illustration 2a) is inadequate for high-fidelity transmission of music, it is wide enough for the faithful reproduction of the voice—yet narrow enough that expensive transmission equipment is not required, thus permitting the use of high-frequency systems.

For transmission over a coaxial cable or radio relay, a set of 12 channels is formed (Illustration 2b), which modulate the carrier frequencies at intervals of 4 KHz (kilohertz) from 64 KHz to 108 KHz. Twelve high-frequency signals are thus obtained, each occupying a band that extends 4 KHz below and 4 KHz above the carrier frequency. However, it is not necessary to transmit the entire band because the two halves of the signal are mirror images of one another, and the complete signal may be reconstructed from only one of the halves. Therefore, one half is filtered out along with the carrier wave (carrier-wave suppression system). The result is a band of 4 KHz, with modulation ranging from 0.3 to 3.4 KHz, leaving an interval of 0.9 KHz separating one channel from the next to avoid the possibility of interference.

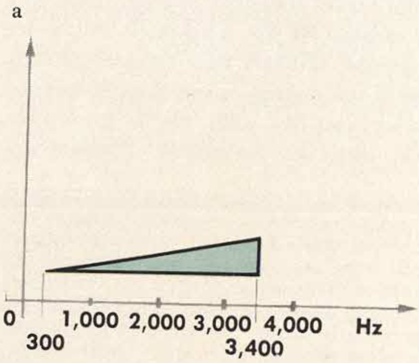
In the channel modulation process, the upper half of the signal is suppressed. The first channel, therefore, has a range up to 64 KHz and remains between 60 and 64. The last channel has a range up to 108 KHz and remains between 104 and 108.

A modulation is then made from the single set of five sets, producing a superset of five channels with carrier frequencies ranging between 420 and 612 KHz and with the two lateral bands 60 KHz apart. (This is because the lowest frequency in the set is exactly 60 KHz.) Here again, the upper band and the carrier wave are suppressed, giving the superset a range between 312 and 552 KHz (Illustration 2c).

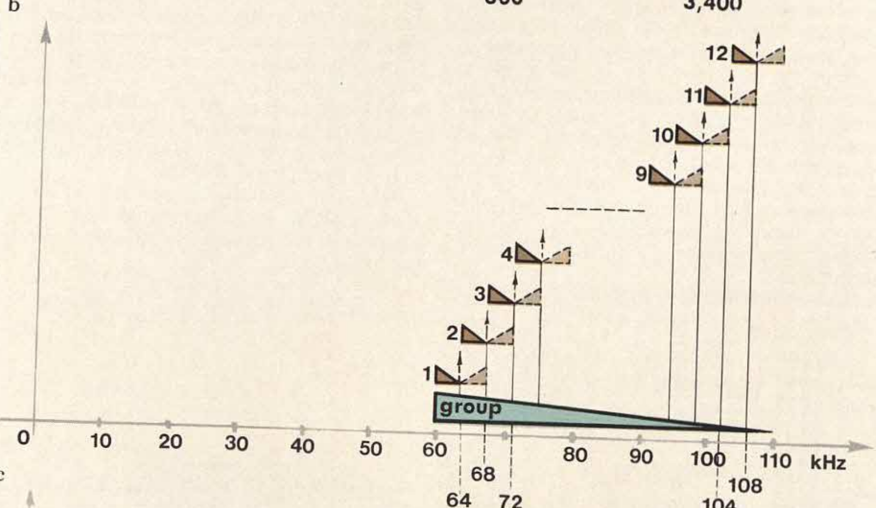
Finally, the superset is modulated to produce 16 supersets. These are similarly modulated and suppressed, except for the lowest set, which is transmitted unaltered from 312 to 552 KHz. The lower band of the second channel (modulated with the carrier frequency of 612 KHz) thus is below the first channel—between 60 and 300 KHz. In this way, a basic band of 960 channels is produced, ranging between 60 and 4,028 KHz (Illustration 2d).

Three of these basic bands may be combined in a coaxial cable. Even more can be combined in a radio-relay system. Both radio relay and coaxial cable allow thousands of communications to be sent simultaneously.

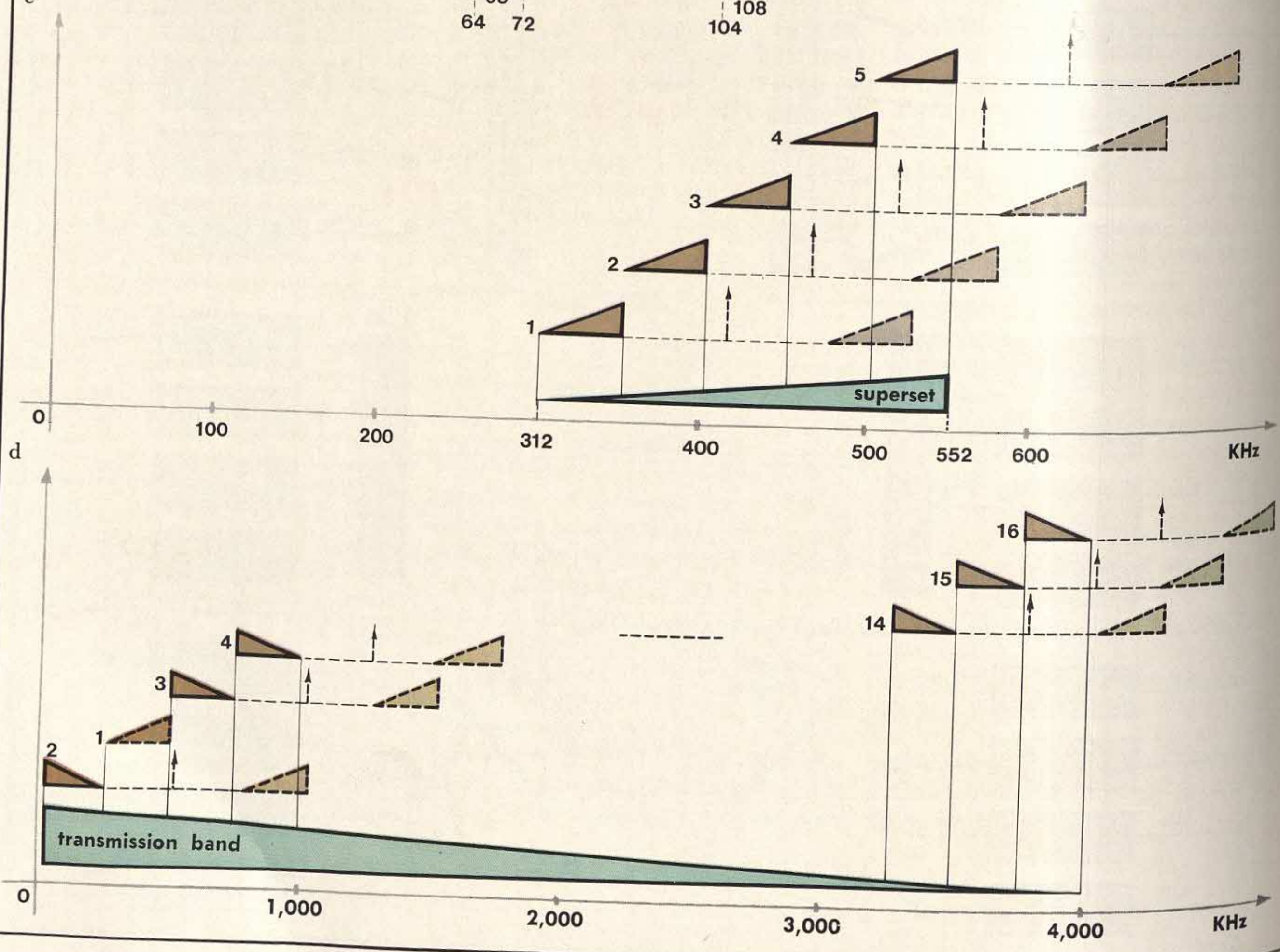
a



b



c



3

THE COAXIAL CABLE—The inner conductor of a coaxial cable is relatively small. It is a solid copper wire that is separated from the outer conductor by plastic insulation. Inner and outer conductors are enclosed in a sheath to protect them from mechanical damage as well as chemical and atmospheric agents.

A coaxial cable also contains metal wires that are used for service communications between transmitting stations and for the remote supply of power and remote control of unmanned, intermediate amplification stations.

COAXIAL CABLE

High frequencies cannot be transmitted along ordinary metallic wire because the resistance within the wire itself causes a large loss of the signal strength within a very short distance. (The capacity of such wire decreases in proportion to the rise in frequency, creating a virtual short-circuit situation.) To transmit such high-frequency signals, a coaxial cable is used. This cable is made up of a conductor placed inside a second, hollow, cylindrical conductor. The outer conductor, insulated from the inner, functions somewhat as a Faraday shield (a device that completely screens whatever it surrounds against static electricity and radio waves). The loss of energy in the outer conductor is relatively small.

A coaxial cable system recently put into operation in the eastern United States uses transistorized repeaters to send 3,600 telephone messages one way through each coaxial pipe. There are as many as 20 such pipes in a coaxial cable. The former system used vacuum-tube repeaters and transmitted only 1,860 telephone messages through each pipe.

MULTIPLEX TRANSMISSION

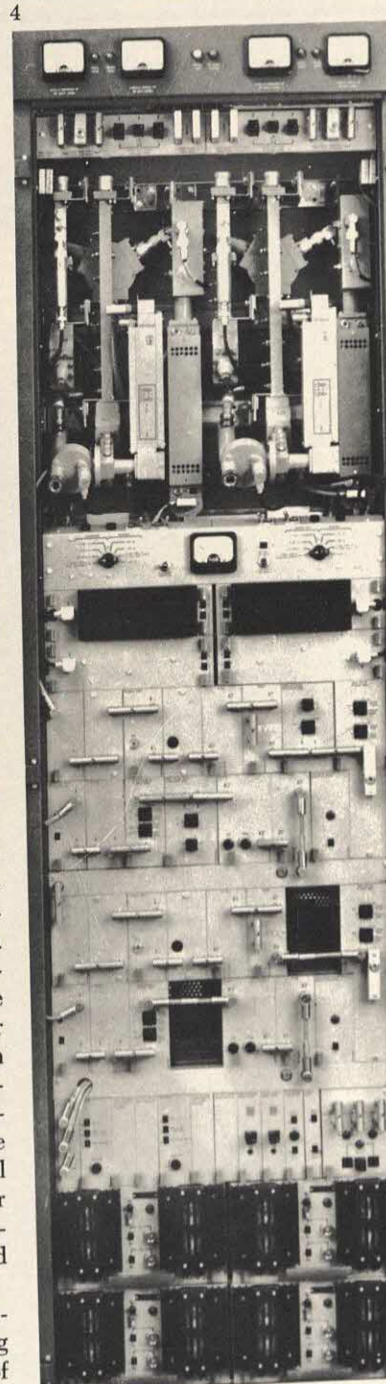
Both radio relay and coaxial cable make it possible to transmit high-frequency signals from one location to another. Both systems can transmit modulated phonic currents, words, or music—just like a radio. However, unlike an ordinary radio that transmits only one message at a time, both radio relay and coaxial cable can carry several messages simultaneously.

With either system, a large number of channels, each with its own frequency and preserving its own identity, are included within the same broad band. Either system may be compared to a train with each car representing a signal carrying its individual message. All the cars travel together on the same track, but their contents remain separated. This is the basis of telephony, which allows sending a number of communications over a given distance, without regard to their origin or destination.

Based on signals given by the telephone caller and the operator, a switching center determines the means of communication of the message. Such a center has two sections. One section selects the receiving center closest to the call's destination and connects the caller with the appropriate line. A transmission section then transfers the phonic currents and signals from the receiving center to the other end of the long-distance line, whether by metallic wire, coaxial cable, or radio relay. Here, another switching center takes the message, locates the telephone of the receiver, and makes the final connection.

Electronic devices such as the transistor may be used to perform switching operations with a speed of the order of a few millionths of a second. Besides being faster than relays and other electromagnetic devices, the electronic devices are capable of performing more complex operations existing at the instant. Solid-state electronics provides the basis for a revolution in both switching and transmission technology.

4



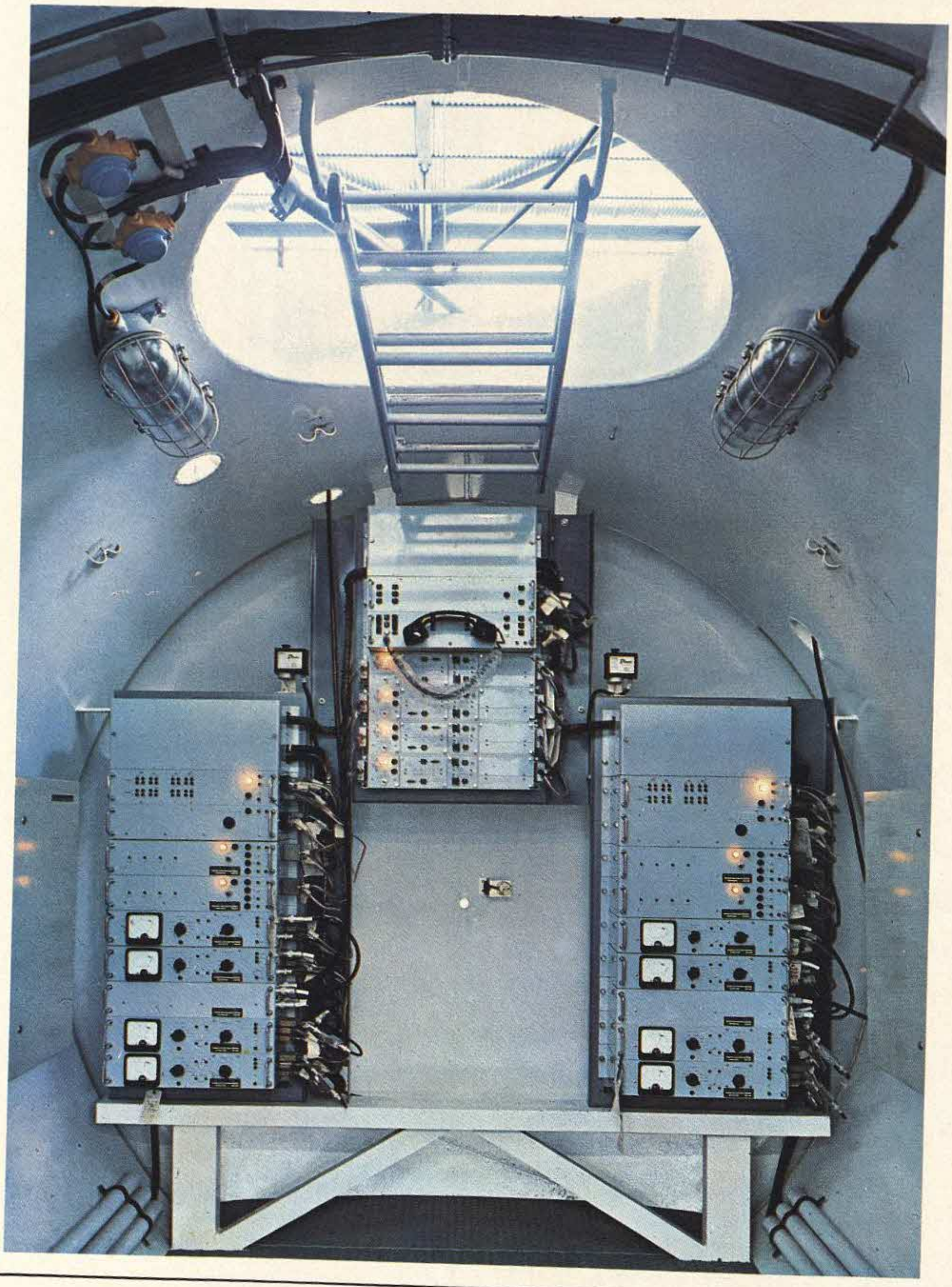
A RECEIVER-TRANSMITTER FOR 960 CHANNELS

5

A RADIO RELAY—A radio-relay system normally depends on several intermediate stations because contact is only possible between points that are in line-of-sight of one another. To increase the distance that may be covered between these stations, they are generally located high—on mountain tops or on towers. Signal repeaters (Illustration 5b) are placed at these points, in view of both the preceding and following stations. The repeater may be passive, merely receiving the signal in one antenna and passing it on through another; or it may be active, with equipment to amplify the weakened signal.

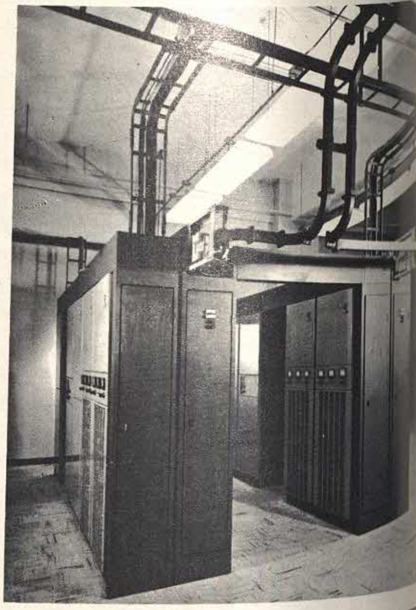
Illustration 5a is a radio relay designed to be buried in the desert, where a tower might not be feasible. Illustration 5c shows a receiver-transmitter station.

a

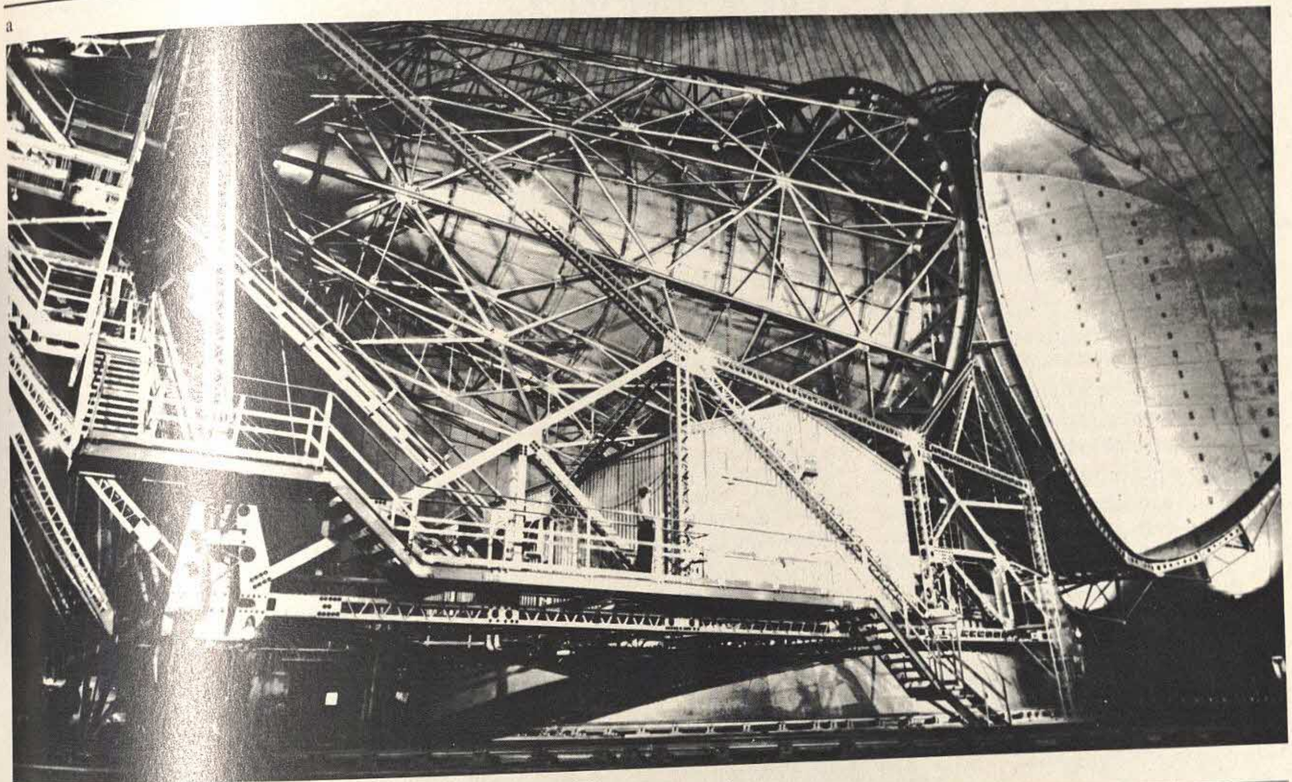


b

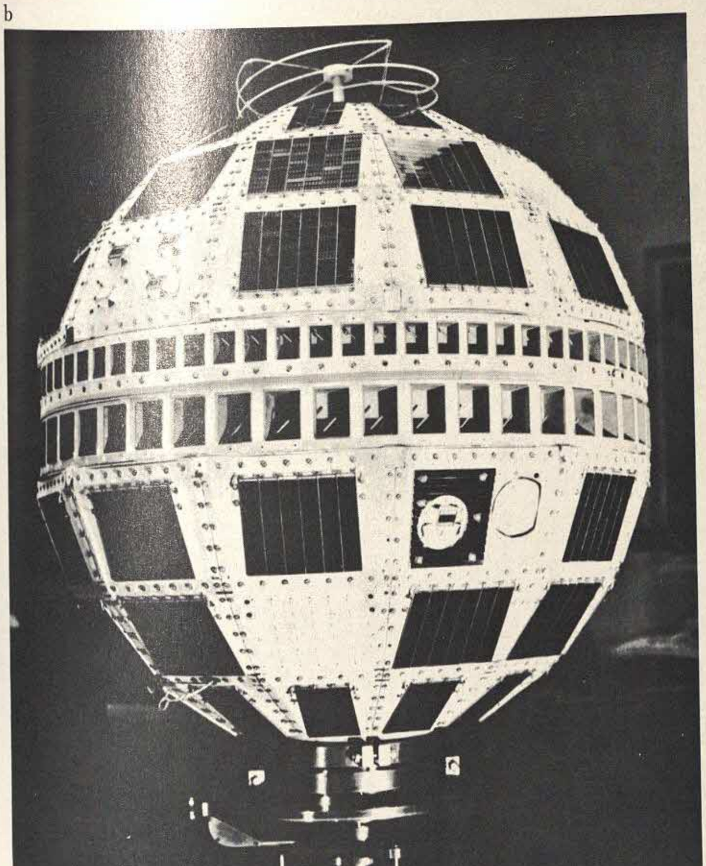
c



a



b



c

SATELLITES FOR TELECOMMUNICATIONS— A stationary artificial satellite may be considered an active repeater in a radio-relay network, receiving signals from an Earth station, amplifying them, and retransmitting them to a station in another hemisphere. Illustration 6a shows a horn-shaped antenna used to transmit signals to, and receive them from, the Telstar satellite (Illustration 6b), which opened a new era in telecommunications. Illustration 6c is the antenna for the Early Bird satellite.

TELESIGNALING AND TELEMETRY | the arrhythmic system

In a remote control system, a central station sends commands to a remote station and receives signals from the remote station concerning its condition. The necessity for this two-way communication requirement is readily apparent. To operate a plant by remote control, an operator must know the exact state of the plant before he transmits commands; moreover, he must also be informed as to what transpires after his commands have been received and executed. For this reason, any command and control system must be equipped with a system for measuring and signaling the results of the commands.

Although the term *signaling system* is used to describe the means by which the central station is informed of the status of its remote stations, in actuality such a system is also a measuring system. Many times, a peripheral station will have more than two states (such as "on" and "off"). For a station with multiple settings, the operator must know the amount of variation required—or existing prior to his command. Such variation can be determined at the control station only if the status of the remote station can be transmitted.

A gas company, as an example, may have one central gas supply. If it is necessary to increase the amount of gas flowing through main A (to satisfy the requirement of one consumer), it is equally necessary to maintain the rate of gas flow in main B (to other consumers). Under such a circumstance, the operator at the central station cannot open completely the supply valve to main A, thereby reducing the amount of gas flowing in main B. If the operator has complete information on the quantities of gas flowing in both mains and if the system permits multiple valve settings, he is able to adjust the valve openings so that main A carries the maximum amount of gas compatible with the requirements of main B. The signaling system must, therefore, be able to transmit more than simple stop-start signals from the remote stations to the central station. It must provide measurements that faithfully represent the existing situation of the whole installation at each and every minute. It must be a measuring system as well as a command system. This does not preclude the transmission of stop-start commands. Such information can be identified with a single electric signal—a pulse normally called a bit. Measurements, however, must be expressed in alphanumeric terms to be of use to the controller, although bits can be used to convey this information.

A sequence of bits—suitably arranged according to the rules of binary mathematics—can be used to express numbers. This is the method most often employed to transmit numbers that represent measurements. It is possible to transmit alphanumeric terms in different ways, the actual choice depending on the constructional considerations and the conditions under which the system is to be used. A single type of system is used in the illustrations. Although the system described is one very commonly employed, other equally effective systems are in use.

The most advanced technical developments of telemetry are utilized in the space program. An unmanned orbiting vehicle is the site of measurements of surface temperatures, magnetic fields, ambient radiation intensity, and other variables necessary in the scientific exploration of space and the development of improved vehicles. In a manned vehicle, additional instrumentation monitors such signals as the astronaut's pulse rate, blood pressure, and respiration rate. Success of the mission and control during the flight require that the results of these measurements be transmitted as frequently as possible to Earth receiving stations.

The telemetry system ordinarily includes: (1) measurement devices; (2) data-processing equipment at or near the point of measurement; (3) electronic communication equipment to transmit the information (by wire or radio, for example) to the desired site; (4) data-processing equipment at the receiving station; and (5) display devices (cathode-ray tubes, loudspeakers, or tape recorders). Items (2) and (4) are the system elements unique to telemetry.

1

THE ARRHYTHMIC SYSTEM—The system illustrated in this diagram is one in which measurements are sent from the remote stations to the central station whenever so commanded; it is also a system in which signals are sent without command whenever a change occurs in the state of the peripheral equipment.

For example, if a "valve closed" signal is in effect it will remain so until a command is received from the operator to open the valve or until the valve is opened by a servomechanism. In either case, the system transmits the proper signal for "valve open" to the central station.

Measurements, however, must be requested by the central station. To determine if the installation is operating correctly, the operator actuates the remote control by pushing the appropriate button. A command goes out from the central station and closes a relay at the remote station, beginning transmission of the requested measurement. The diagram indicates the equipment installed at the remote station, including an analogue-to-digital converter that generates an electric signal (manometer with potentiometer). When the relay is closed (as a result of the command received), the electric signals of the potentiometer are sent to the analogue-digital converter that converts the voltage signals into pulses. The number of these pulses is now proportional to the voltage value of the signal—for example, 10 volts = 1,000 pulses and 5.55 volts = 555 pulses.

These pulses are then sent to a register, consisting of 10 binary circuits connected in series—an arrangement capable of counting up to 1,024 pulses. Conducting these binary circuits in parallel yields an output that permits a 10-bit representation of the pulses generated by the analogue-digital converter—a method that also expresses the voltage output of the potentiometer.

The position of the potentiometer corresponds to the pressure in the pipe (or, as in the example given in the text, the amount of gas in the main).

These signals are received at the central station by a register that transmits them to a digital-analogue converter, where the measurements are transformed back into units easily recorded by standard instruments. The complete chain of events at the central station—between the request for a measurement and the reading of that measurement—is described in the following paragraphs.

The operator at the central station presses a button corresponding to the measurement he requires. The command (for example, 135) goes out through the remote control system to the remote station. The relay corresponding to 135 closes at the remote station. At the same time, an indicator light goes on alongside the button pushed (showing that 135 has been requested) and another indicator light goes on at the instrument where the measurement will be read. In this way, the control panel always displays the last command sent out.

The relay actuator at the remote station closes contact C_1, sending the potentiometer voltage through to the analogue-digital converter. The converter emits a train of pulses proportional to this input signal. A binary counter receives these pulses. Meanwhile, a pulse is sent from contact C_2 of the actuator relay, through the delay circuit R_1, to both the 10 AND gates (A through J) and the OR gate A_1. This pulse has the effect of transferring the contents of the binary counter to the translation register and sending the start pulse to the transmission cable. The oscillator controlling the movement of the sliding register is unblocked by the delay circuit R_2, permitting the

translation register to transmit the coded measurement. The binary circuit BP_1 ascertains the parity of the message after counting the number of output pulses from the register, it adds an additional pulse in the twelfth position if this is required to make the total number of pulses even. The delay circuit R_3 serves the purpose of inserting the state of the parity binary circuit in the twelfth position of the message.

Upon arrival at the central station, the pulses enter a translation register and are transferred only if the parity conditions are correct. The binary circuit BP_2 is inserted to check that the number of pulses received is even. The start pulse (acting through F_1) brings the register back to zero before the arrival of the other pulses. The sliding register transfers the message code to the digital-analogue converter. The converter generates a current that is sent to the measurement-reading instrument, completing the operating cycle.

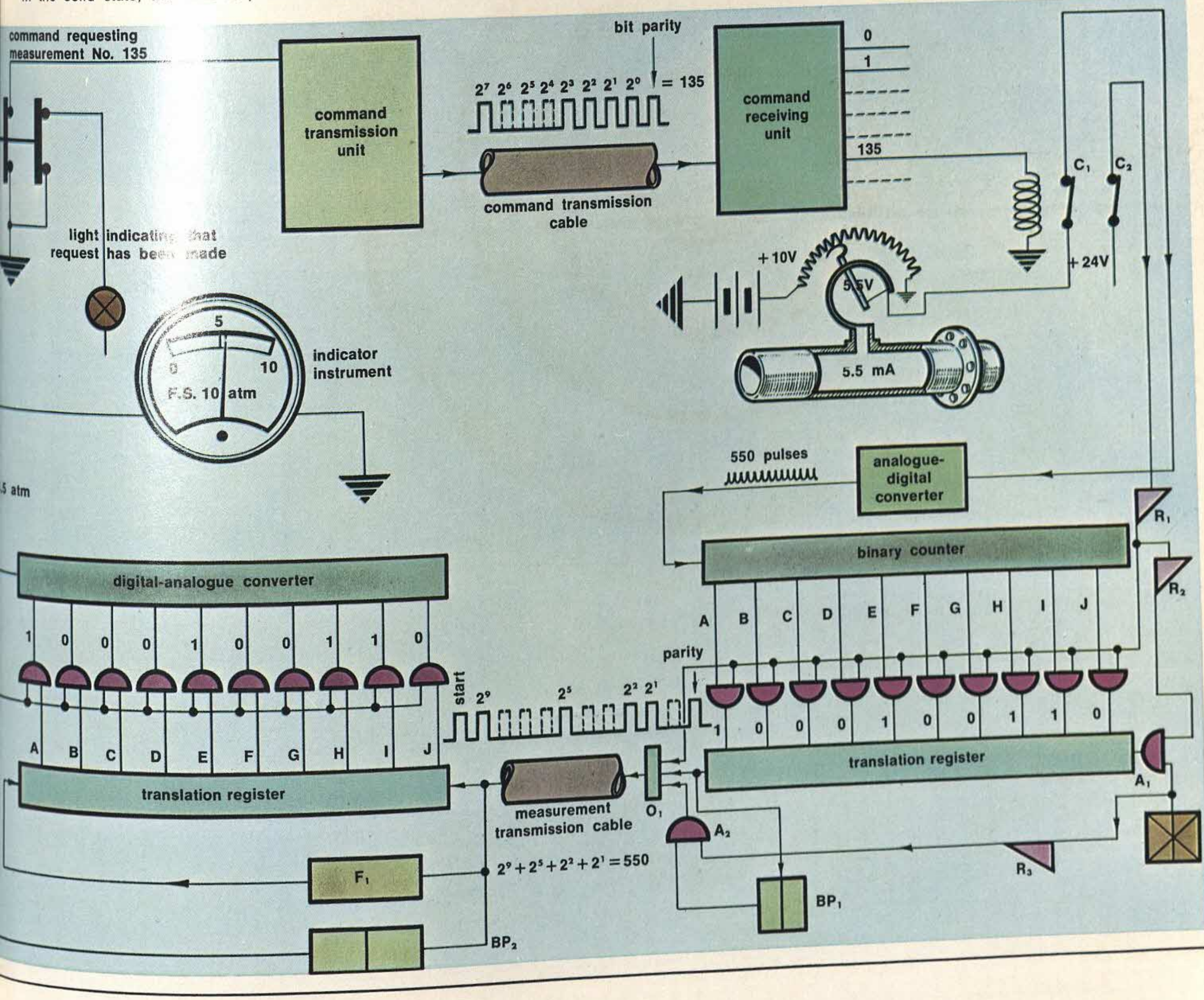
The time required to complete the operation described is about 1 second in slow systems and 0.2 second in the faster systems. This time is relatively long when compared with the relay actuators, which normally require only 50 milliseconds.

If the whole of the system described were in the solid state, the time required for the whole operation would amount to no more than a few milliseconds.

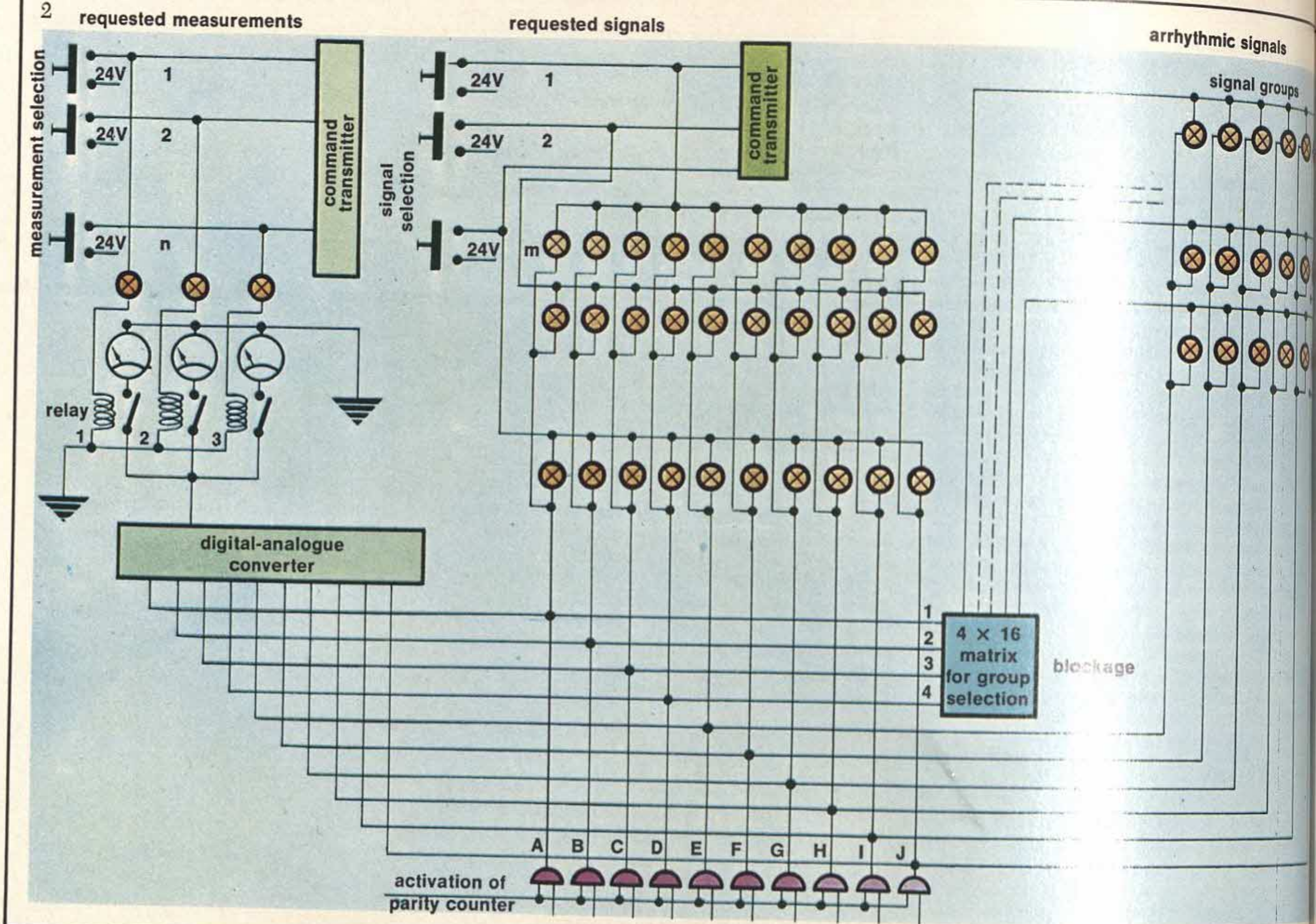
The arrhythmic method of obtaining measurements is also used for signals. Signals are generally divided into two distinct groups: those transmitted at the request of the operator; and those automatically transmitted from the remote stations. The former are usually requested by the operator to update the situation regarding variables that are questionable at the central station. Automatically generated signals, however, are of the utmost importance. They must be sent to the central station as soon as changes occur because they are generally critical and commonly vital to the safety of the plant. Alarms are a good example of this type of signal.

The need for these two different types of signals dictates different requirements within the system. Both types, however, are transmitted from the remote station as part of a byte. Referring again to the example in the illustration, the length of the telegraphic message (and, in a certain sense, the system itself) remain unchanged if signals are requested in bytes consisting of 10 bits. In calling for a byte, the operator is aware that it will arrive in a package of 12 bits (including start and parity) and that the message will consist of state variables (for example, open-closed, stop-start, and so forth), not measurements. The first bit output (in the example shown in the illustration this would be from AND gate **A**) will correspond to the first state variable of a certain group, indicating that a certain state of a component exists (or does not exist). If the operator understands that the signals inserted in the 10 AND gates (**A** through **J**) of the remote station can be extracted from the corresponding AND gates of the central station and sent to the signal lights after codification, he will know how to read those signals.

The same discussion applies to automatic response signals, which always pass through the same AND gate in both stations. However, because these signals have not been requested by the operator, they can only be identified if they contain some further information indicating their nature. The first bits of the message are utilized to achieve this identification. In the example, the first four bits are used for identification purposes; these initial bits direct the information to the correct display indicator. Automatic response signals, therefore, are always preceded by a code. Like a key, this code opens a path along which the remaining bits proceed to a single, defined destination. In the example, the first four bits (**A, B, C,** and **D**) constitute the selection codes and the remaining six bits (**E** through **J**) the information.

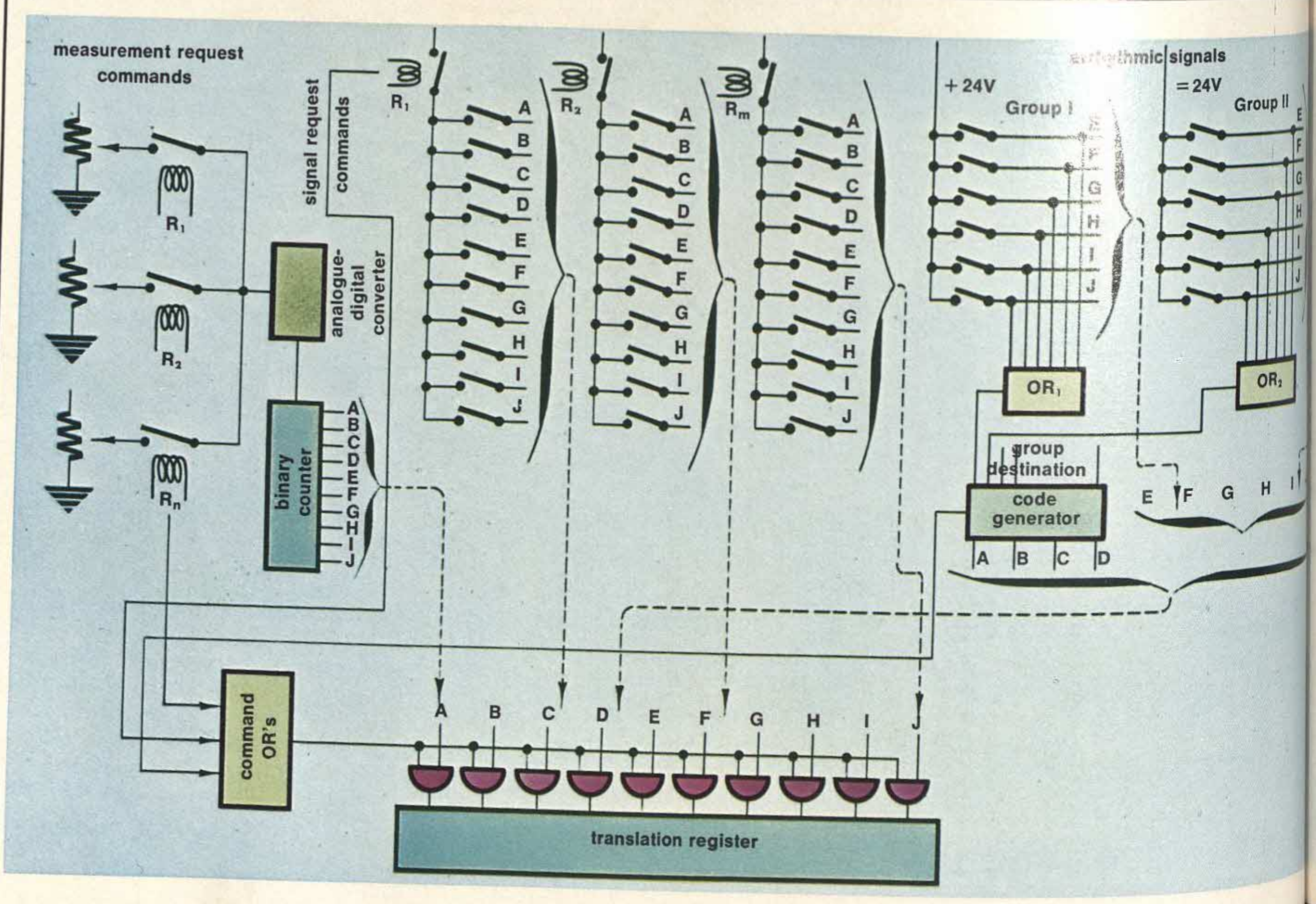


2


requested measurements
measurement selection
24V
1
2
n
command transmitter
relay
digital-analogue converter
requested signals
signal selection
m
4 × 16 matrix for group selection
blockage
A B C D E F G H I J
activation of parity counter
arrhythmic signals
signal groups

3

a


measurement request commands
R_1
R_2
R_n
signal request commands
analogue-digital converter
binary counter
A B C D E F G H I J
R_m
arrhythmic signals
+24V
Group I
=24V
Group II
OR_1
OR_2
group destination
code generator
command OR's
translation register

CENTRAL STATION FOR THE RECEPTION OF REMOTE CONTROL SIGNALS AND MEASUREMENTS—This diagram shows the three parts of the remote control systems.

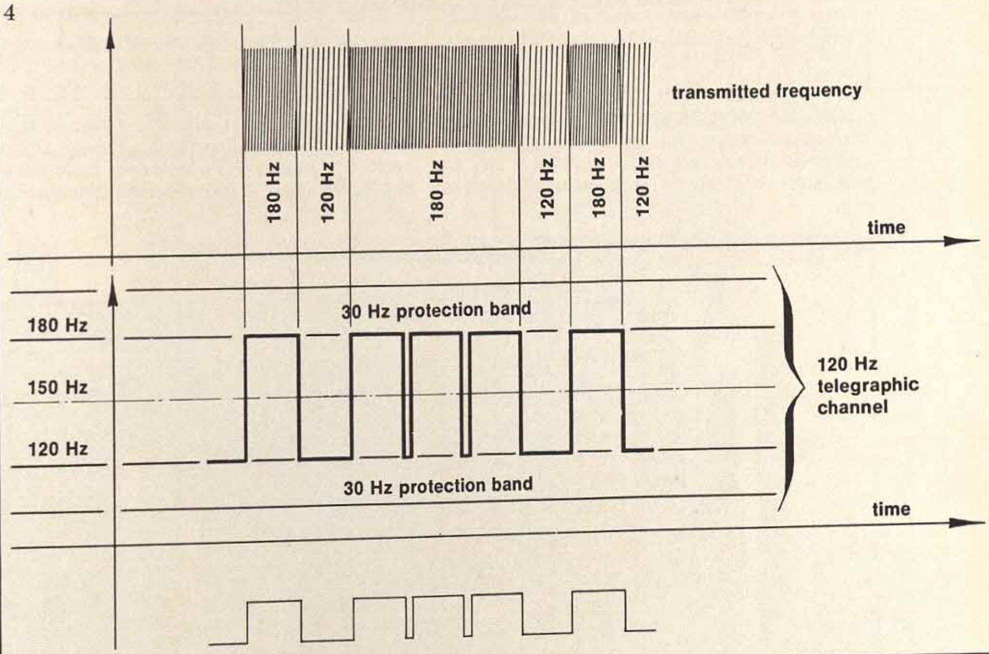
The first part (on the left) involves requested measurements. Pressing the **n** button activates the corresponding measuring instrument through the **n** relay. The requested measurement arrives at the central station in codified form through the AND gates (**A** through **J**). It then goes to the digital-analogue converter, to the signal light corresponding to the measurement, as well as to the group-selection matrix for the automatic response signals. Only one of the instruments is activated by the push button. The other signal indicators—inasmuch as they are not activated—do not receive any signals; the indicator lights of the automatic response signals remain inactivated because of the blocking control incorporated into the matrix. The blocking command is automatically transmitted to these indicators whenever a measurement or state variable signal is requested.

The second part (in the center) involves requested state variable signals. In such a request only the row of indicator lights relating to the requested indication is activated by the push button. The blocking signal is automatically transmitted.

The third part (on the right) is concerned with automatic response signals, which are generated without a command from the central station. The code preceding the information activates a row of indicator lights corresponding to the response group (this is done through the matrix). A four-bit code permits the activation of 16 different groups through the different combinations possible within the four code bits. The code is followed by information bits (in this case, six).

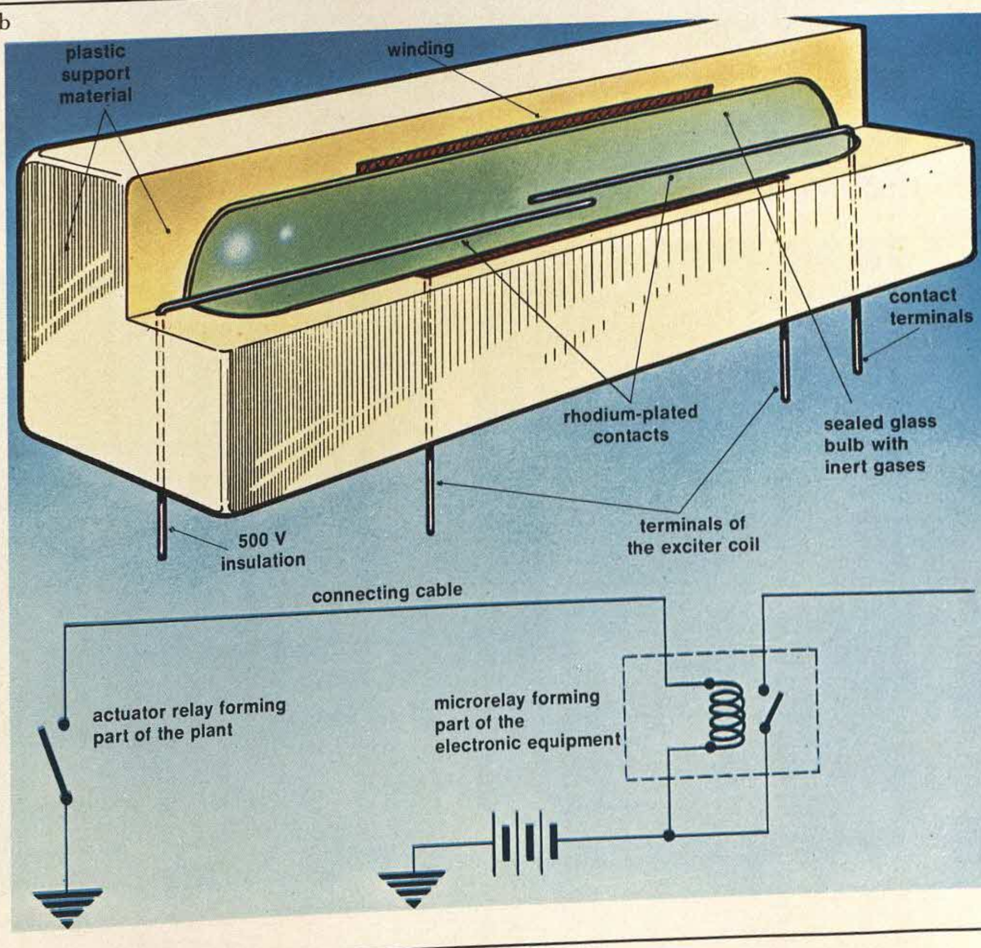
TELEGRAPHIC BANDS—The arrhythmic signaling system has been portrayed in general terms in the preceding illustrations. Only the basic layouts have been shown. The examples provided heretofore have shown that the message has been entrusted to the test of simple parity only. This is not sufficient in practice, however. Messages are transmitted with a series of pulses only in some rare cases. Preference is shown for transforming the bits into signals of a specific frequency—for example, by using a frequency of 180 Hz to represent an "exists" bit (state 1) and by using a frequency of 120 Hz to represent a "does not exist" bit (state 2). Such a system makes it possible to transmit several messages simultaneously through the same cable by translating the messages from various stations into different frequencies. Each transmitting station, therefore, has a well-defined frequency band at its disposal. These bands (telegraphic channels) generally have a band width of about 120 Hz for messages transmitted at a rate of 50 bits per second.

4



THE REMOTE STATION—Illustration 3a shows the remote station of a remote control system reacting to commands from the central station and also providing automatic state variable response signals. The measurement-request equipment is similar to that shown in Illustration 1; it is shown here to complete the block diagram of a remote station. The signals requested are provided in groups of 10 after the command received from the central station has activated the whole of the group. This activation is begun through relay $\mathbf{R_n}$, which closes the **n** bar that feeds the actuator relays of the **n** group of signals. These signals are transmitted in the same way as measurements—that is, they are sent through the AND gates to the translation register (shown in Illustration 1). The start and parity bits are generated in the same way as they are in measurements. Bytes of six bits each are sent out with their data every time one (or more) of the signals in the group closes its actuator relay. The codes (the first four bits that determine the destination of a signal) are transmitted through the OR_1—OR_{16} gates. At the same time, a command is sent out to activate the transmission of both the destination code bits (**A** through **D**) and the information bits (**E** through **J**). A schematic of this signal grouping is shown in Illustration 3a; actually, the signal relays are spread out over a vast area and can also be interconnected. The signal relays are redundant—a relay incorporated into the plant itself controls a miniaturized relay (Illustration 3b) included in the electronic system. The electronic system is generally designed to eliminate electrical disturbances (with a value of several hundred volts) that may interfere with the electronic equipment. This redundant system destroys continuity between the actuator and repeater relays.

b



TRANSFORMERS | low voltage–high voltage–low voltage

A transformer is a static device; it does not have any moving parts. Moreover, it neither receives nor supplies any mechanical power, but is concerned solely with electric power. The extensive development of alternating-current machines and the predominance that this type of current has assumed in industrial applications are due to the possibility of transforming low voltages into high voltages, and vice versa. The transformer is of enormous practical importance. Most electricity is not produced close to the places where it is used. In many cases, electric power stations are situated in mountains where they can utilize the water of large storage basins. On the other hand, if the power station is a thermal one, it will generally be located where fuel is available to produce steam, or gas for turbines; or, it may be close to a port where fuel supplies can be easily obtained by marine transport. Energy production centers (excluding nuclear

1

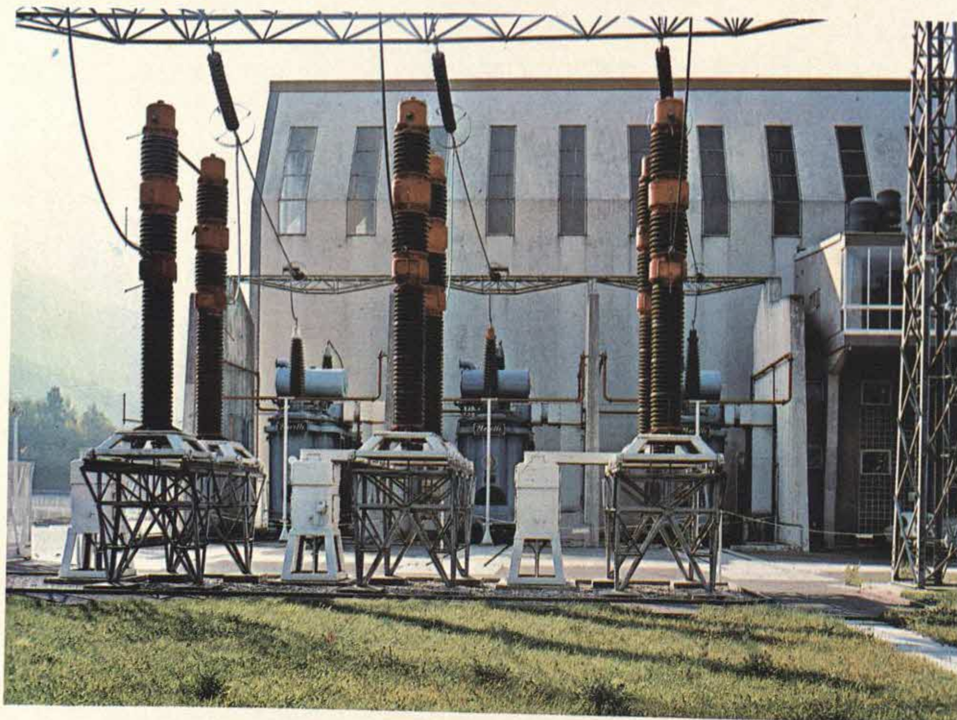
ENERGY TRANSMISSION—The long-distance transmission of the enormous quantity of electric energy that is required each day by industry and private consumers is achieved by means of a dense network of high-voltage transmission lines. Raising the voltage of the current by means of a transformer permits the generated power to be conducted to the transmission line at a considerable current intensity. This makes it possible to construct the transmission lines with conductors that have smaller cross-sectional areas. The metal needed for the construction of the cables is very costly. Reducing the quantity of this expensive metal in long power lines is an obvious economic advantage. In addition, the lighter transmission lines need less expensive towers to support them and the dissipation of power along the conductor (a resultant of the Joule effect) is reduced.

The electric power station pictured in Illustrations 1a and 1c is located in the Alps. It produces electric energy by exploiting a head of water of about 1,000 meters. This station is equiped with Pelton hydraulic turbines, each having a power rating of 110,000 kilowatts. The voltage of the current produced by the alternators is raised to transmission values by transformers (Illustration 1b) installed at the start of the transmission line. Very high transmission voltage (over 400,000 volts) can be attained. At the end of the transmission line there is a substation where the voltage is lowered again.

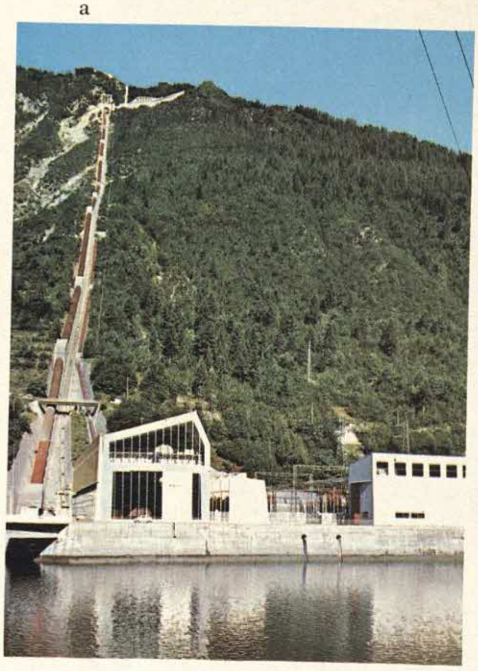
The distribution voltage is not yet the voltage at which the current is to be used. It still has a high value and may vary from 500 to 3,000 V.

A further voltage reduction is achieved by means of smaller transformers installed at suitable locations. Frequently (especially in the countryside) transformers are installed on the poles of the distribution line to serve small inhabited centers such as workshops and construction sites (Illustration 1d).

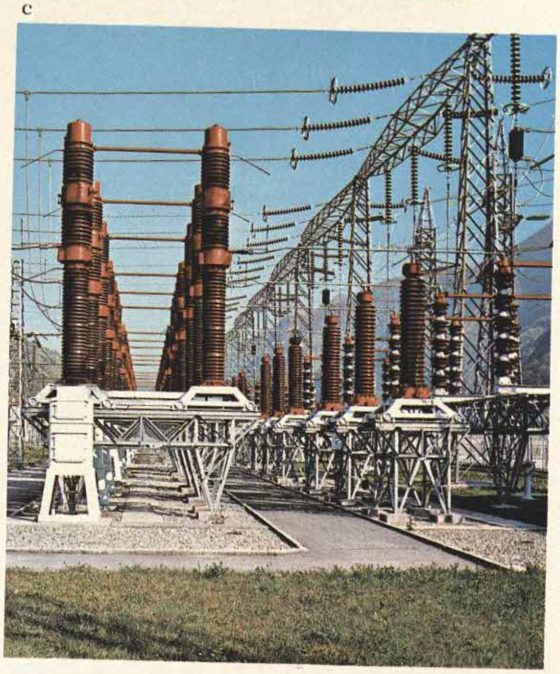
b



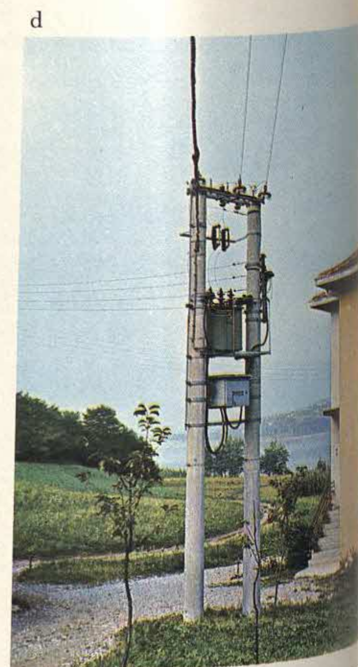
a



c



d

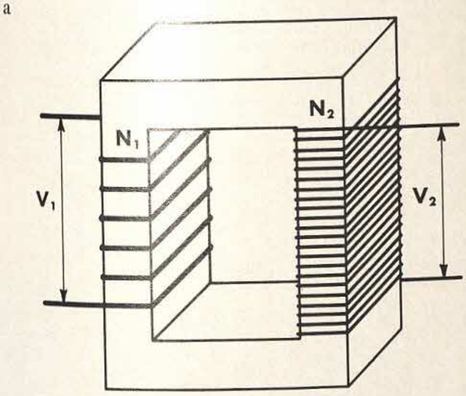


2

THE PRINCIPLE ON WHICH A TRANSFORMER FUNCTIONS—A transformer can be either single-phase or three-phase. (Other types of polyphase transformers have had no extensive applications.) The principle on which the single-phase transformer is based is described below. The three-phase transformer (Illustration 2b) is derived from this principle.

A transformer consists of a magnetic core made of highly permeable material. For the sake of constructional simplicity, this core is arranged as shown in Illustration 2a. This core has to contain an alternating magnetic flux and is made from a series of steel laminations, insulated from one another and held together in an appropriate manner. This solution is adopted also for the construction of the cores of other electrical machines and has the purpose of reducing the hysteresis currents to a minimum. The core carries two fixed windings consisting of tightly wound turns; the two windings are insulated and separate from each other. When an alternating voltage $\mathbf{V_1}$ is applied to the terminals of the input winding, a transformed voltage $\mathbf{V_2}$ will manifest itself at the terminals of the output winding. The ratio between these two voltages is known as "the transformation ratio of the transformer" and is almost equal to the ratio between the numbers $\mathbf{N_1}$ and $\mathbf{N_2}$ of the turns in the two windings. The first $\mathbf{N_1}$, which is supplied with the voltage $\mathbf{V_1}$, is called the primary winding. The other $\mathbf{N_2}$ is known as the secondary winding. However, this definition is not a rigorous one: a transformer is reversible. Either of the two windings can fulfill the function of either the primary or the secondary winding, according to the manner in which the transformer is being used. From the constructional point of view, the two windings are classified as low-voltage winding $\mathbf{N_1}$ (with the smaller number of turns) and high-voltage winding $\mathbf{N_2}$ (with the larger number of turns).

a



The transformer is called a step-up transformer when the low-voltage winding acts as the primary winding; it is known as a step-down transformer when the function of the primary winding is fulfilled by the high-voltage winding.

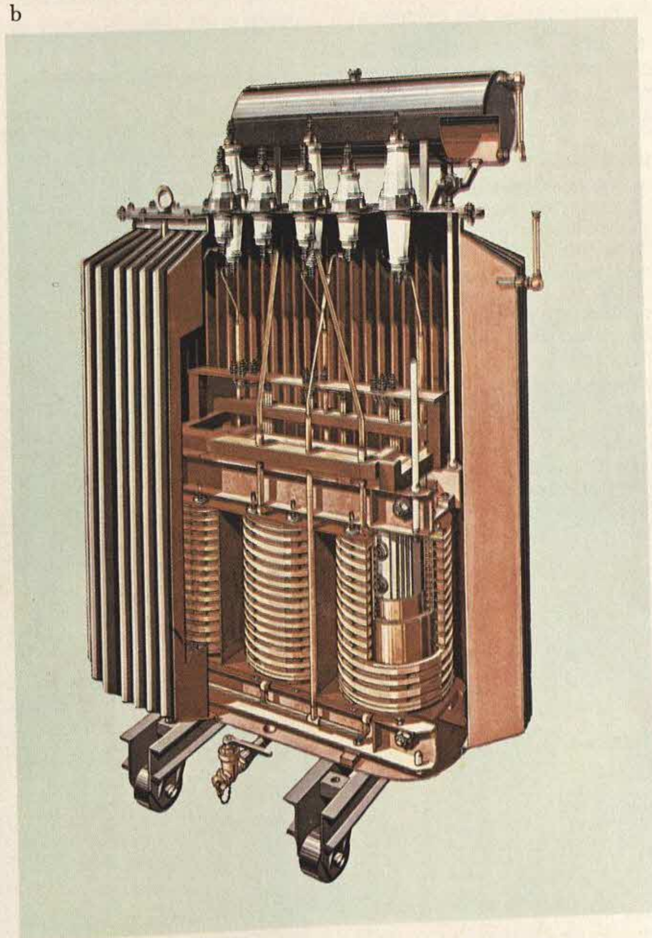
This principle may be further clarified by considering an idealized transformer, one in which all the losses in the iron and the electric resistances of the windings have been reduced to zero and in which there is no dispersion of the magnetic flux—a transformer in which the whole of the magnetic flux remains inside the core without any dispersion. Such a transformer would be a perfect one.

The conditions under which the transformer functions are known as no-load (when the secondary circuit is open) and under load (when it is connected to a load).

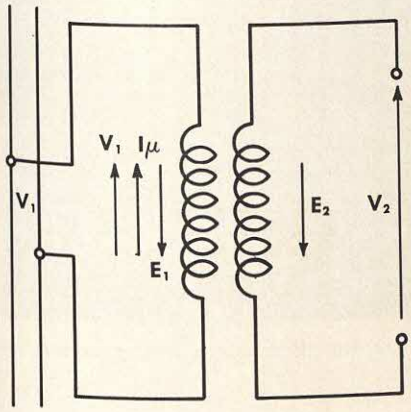
When a transformer operates under no-load conditions (Illustration 2c), a voltage $\mathbf{V_1}$ is applied to the primary voltage. This produces and maintains a current $\mathbf{I_\mu}$ in this winding. This current—known as the magnetization current—generates a flux in the core, inducing in the primary winding an electromagnetic force $\mathbf{E_1}$ equal and opposite to the applied voltage. The same flux induces in the secondary winding an electromagnetic force $\mathbf{E_2}$ that provides the voltage $\mathbf{V_2}$ across the secondary terminals.

When a transformer operates under load (Illustration 2d), a load circuit is applied to the secondary winding. The secondary electromotive force $\mathbf{E_2}$ causes a current $\mathbf{I_2}$ to flow through the circuit. This secondary current—flowing through the secondary winding of the transformer—disturbs the equilibrium of the magnetic flux in the core. Because the electromotive force $\mathbf{E_1}$—and therefore the voltage $\mathbf{V_1}$—have to be maintained constant, a new current will instantaneously occur in the primary winding. This current is necessary to re-establish the equilibrium and restore the initial value of the flux. The new current is known as the primary reaction current. It is added to the pre-existing current $\mathbf{I_\mu}$ and gives rise to the primary current $\mathbf{I_1}$. The secondary and the primary reaction currents will vary as the load varies, but $\mathbf{I_\mu}$ remains constant. However, because its value is very small, the primary current $\mathbf{I_1}$ will differ only slightly from the primary reaction current, and the ratio between $\mathbf{I_1}$ and $\mathbf{I_2}$ is equal to the inverse ratio between $\mathbf{N_2}$ and $\mathbf{N_1}$ (Illustration 2a).

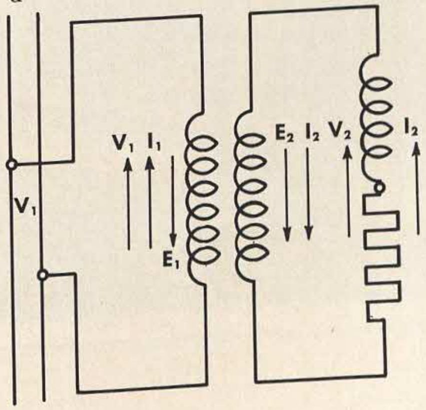
b



c



d



power stations) are therefore found most often in areas away from the centers of energy utilization—sometimes at considerable distances.

Another consideration is that modern industrial nations must be covered by an electrical network sufficiently interconnected to ensure that key areas will not suffer from the failure of any one power station. This creates a number of economic problems. If the machines that produce energy are to utilize fully the thermal or mechanical power they receive, they must supply the current at a voltage compatible with the efficient operation of their equipment. For reasons of safety, as well as of construction, the voltage supplied to equipment using the electric energy produced must be kept within a few hundred volts. There is an immense amount of electric power required, and the demand for it increases continually. Under these circumstances, efficiency in the production, transportation, and utilization of electric energy is mandatory.

The basic problem, therefore, is transporting the electric power (produced by the most efficient means) across the distances required, and delivering it to the users in the most suitable form at a minimum of cost.

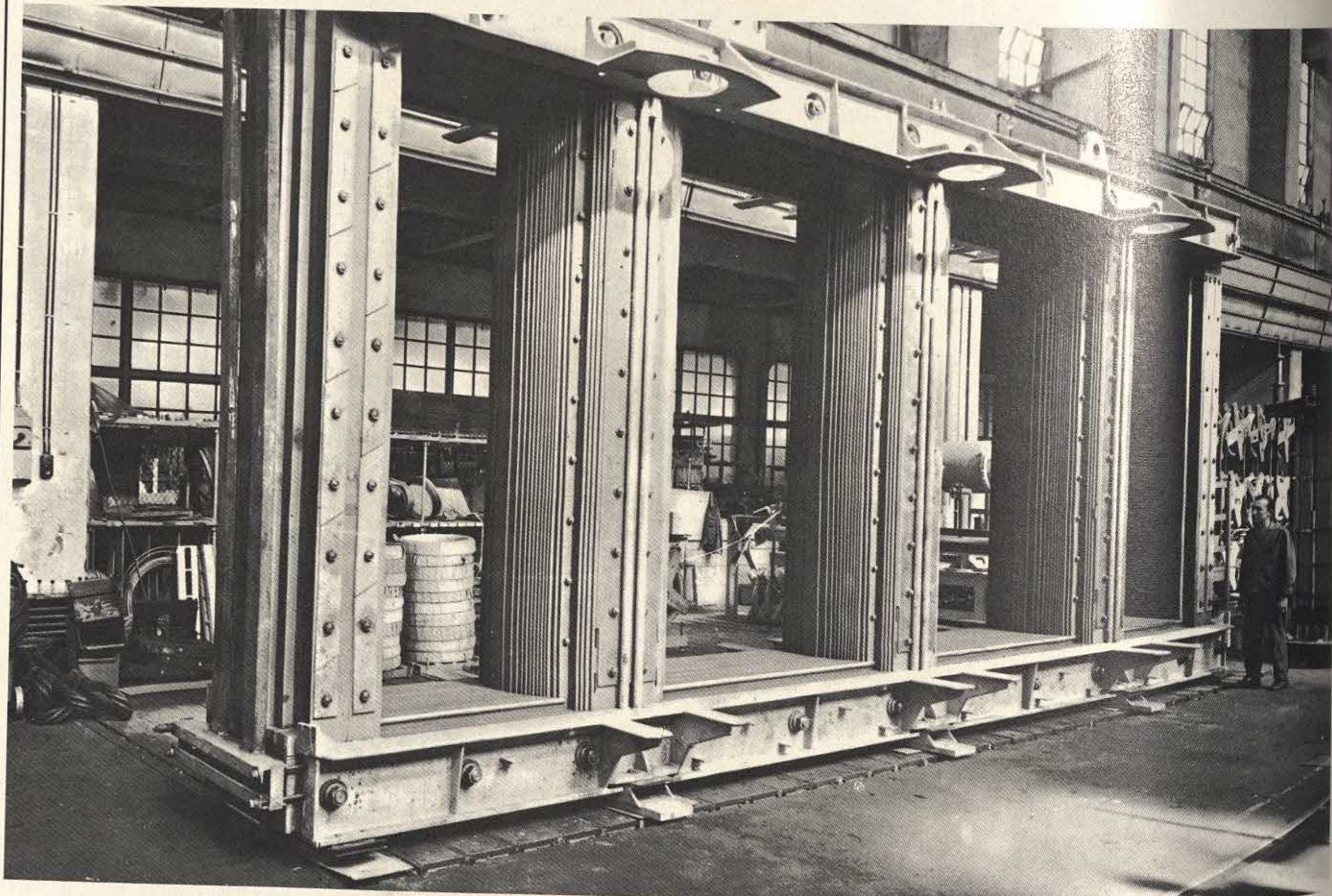
This problem is complicated by a restriction: for both economic and structural reasons, the size (in cross section) of transmission lines is limited. The temperature of a conductor increases when an electric current flows through it. This means that a certain amount of power (in proportion to the current) is lost in heat. More precisely, the power dissipated in heat is proportional to the square of the

3

THE CORE OF A LARGE THREE-PHASE TRANSFORMER—Some of the constructional details of the core of a large three-phase transformer are shown in Illustration 3.

The core of a transformer is a very important element. Apart from having to contain the magnetic flux produced by the windings, it also must constitute a strong supporting element for these windings. Therefore, as shown in Illustration 3a, the core is an imposing and massive structure. The portion that constitutes the three central columns and the transverse members that join them together requires very delicate and accurate construction. It is not constructed in a single piece, or even in pieces of considerable dimensions, but is formed by the union of an extremely large number of laminas made from a special steel. This constructional solution has been adopted to reduce hysteresis currents. In fact, if the nucleus were homogeneous and made of one piece, the effect of the magnetic field would be that of generating induced currents that would be distributed in a complicated and irregular manner inside the mass of the metal. These currents, which are also known as Foucault currents, are harmful because they constitute a dissipation of the power in the form of heat (the Joule effect). The core—formed by an assembly of laminas separated from each other by a thin layer of insulating materials—has a very high average resistance, much greater than it would have if it were made of one piece. The average intensity of hysteresis currents—and the power that is absorbed by them—is, therefore, very small.

a



intensity of the current. Therefore, if the conductor is to minimize the heat that is produced, while permitting the passage of a current of high intensity, it must have a large cross-sectional area. In order to permit the transmission of a specific amount of power, therefore, the voltage must be raised and, consequently, the current must be diminished. In fact, electric power, voltage, and current are linked by the relationship $W = EI$ (where W is watts, E volts, and I amperes), in conformity with Ohm's law.

All of these problems are solved by the application of the transformer. The generator, in fact, can supply the required power at values of the voltage and the current that are best suited to its own operational characteristics. For the purposes of transmission, the voltage is then raised at the power station where the transmission line starts, in order to avoid excessive heating of the wires (which results in a power loss). The wires, in turn, can have small cross-sectional areas and this will make them cheaper and easier to install. In order to achieve the long-distance transmission of power on the order of many thousands of kilowatts, it has been necessary to adopt increasingly higher generation voltages, transmission voltages, and distribution voltages. In each voltage transformation, the current will transform inversely to the voltage in accordance with the formula described above; the power cannot change apart from the inevitable power losses that occur during transmission and transformation.

b

NETWORK SYSTEMS

In metropolitan areas, forms of power distribution systems known as network systems have been developed. Begun in New York City in 1922, the secondary network system provides that the secondaries of the final transformers, instead of feeding only a few customers, are all connected directly together over a wide area by large conductors to form a network, and all users are fed from these secondary lines running along each street. The transformers, instead of being small units, are larger—up to 500 kw or more—and are frequently fed directly from the generating stations at generator voltage, often 13,800 V, by underground cables. The transformers are located in underground vaults, usually under the streets.

Network protectors—special circuit breakers connecting the transformers to the secondary lines—open whenever power feeds from the network to the transformer, even in small amounts. This may occur when the primary circuits are opened at the generating stations and the transformer idling losses are supplied from the network, or when a defect occurs in a transformer. Otherwise, whenever power can feed from the transformer to the network, the protector closes. Short circuits in the low-voltage network, which rarely occur because of the large conductors, well insulated and underground, are cleared by the melting of the conductors in the region of the short circuit or by the melting of reduced sections known as limiters. Since every line is fed by at least two transformer banks, one at each end, the first melting of the cable does not cause any interruption in service to the consumer. Special modifications are used in industrial plants and large office buildings. This system is economical in regions of high load density.

TRANSISTORS—I | replacing the vacuum tube

The advent of the transistor revolutionized the field of electronics and greatly expanded its horizons. The transistor has many characteristics that make it far superior to the triode, an equivalent electronic structure for which it often substitutes. The triode and the transistor are the two most important active elements in an electronic circuit. Active elements are those that introduce electric charges into a circuit and can act as amplifiers.

An element of an electric circuit is sometimes depicted as a box with wires entering and leaving. The wires carry an electric current that can be determined for each specific voltage and resistance. An analysis of the elements inside this closed box permits the determination of how much of the current entering the box will leave it.

Among the components of a closed circuit box are the passive elements, such as resistors, inductors, and capacitors. These elements have two terminals, each indistinguishable from the other, that can be connected in a variety of ways without introducing variations in the circuit. In a special category are the diodes. They act like two-way resistors: in one direction, they contain a very high resistance; in the other direction, their resistance is weak. Finally, triodes and transistors are active elements because they can modulate the currents passing through them.

The triode consumes more energy than does a transistor because a large amount of power is dissipated in the form of heat and is hence lost as useful output. The transistor, however, possesses its own electric charges whose action depends on the input and regulation voltages. Thus, in a transistor, the energy consumed is a function only of the output.

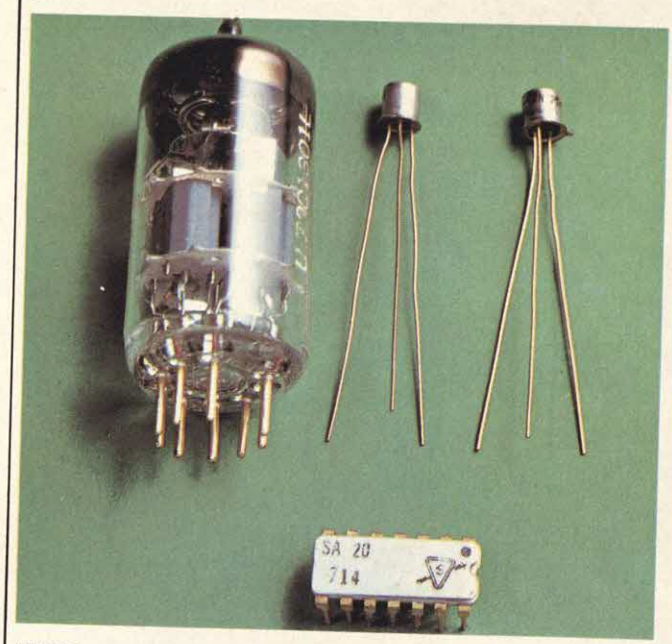
ELECTRONIC CHARACTERISTICS

Because they consume little energy, transistors can be used in instruments where the energy is provided by batteries. Such instruments are used in astronautics, in telephony, in telegraphy, and also in intercontinental television transmission by satellite. Transistors can be of very small size even though they must furnish large amounts of power. For example, a transistor that must put out several watts has a volume no greater than one or two cubic centimeters. Most of this volume is taken up by its container, not by the transistor itself. This excess volume is used to house a relatively large metallic body. The transistor is mounted in contact with the body, called a heat sink, and thus the transistor can dissipate what little heat it generates (such heat can damage the transistor). The small dimensions of the transistor can be further reduced, permitting the construction of certain types of apparatus that, at first, were unthinkable.

Vacuum tubes used large amounts of power—sometimes hundreds of volts—as

1

a



b

THREE GENERATIONS OF ELECTRONIC COMPONENTS — The first generation corresponds to vacuum tubes, the second refers to electronic transistors, and the third to monolithic integrated circuits. This evolutionary development has resulted in an ever-increasing reduction in circuit volume, in the dissipation of electrical energy, and in a resultant reduction in heating.

Illustration 1a shows a double triode, for amplification states up to 150 MHz and a power of 0.5 W. To the right are two tran-

well as precise construction. Transistors, however, can operate with the much smaller amounts of power provided by batteries. Because the transistor normally uses only a small amount of power, and uses that only when it is actually functioning, batteries last a relatively long time. On the other hand, transistors cannot function at high frequencies (for example, 10,000 megacycles per sec) as vacuum tubes can. It is for this reason that in radar bands, vacuum tubes are still used to generate microwaves. Research is under way to adapt transistors to higher frequencies.

Another difficulty is that transistors sometimes produce a hum. This hum (or white noise) is produced by an alternating current superimposed on that provided by the transistor and transformed into sound by the loudspeaker. The source of this noise is in the movement of electrical charges within the transistor and is primarily a function of its temperature.

THERMAL CHARACTERISTICS OF TRANSISTORS

Although vacuum tubes can be built to withstand temperatures of hundreds of degrees, transistors cannot. If they are made of germanium, transistors function best in temperatures less than 35 to 40° C (95 to 104° F). Those made of silicon can function at much higher temperatures, but less than 100° C (212° F). This limitation applies to environmental temperature, the reference point used when determining the cooling of the component. If the environmental temperature rises, the cooling of the transistor (whether by natural or forced ventilation) lessens and the temperature of the transistor rises above the permitted limits.

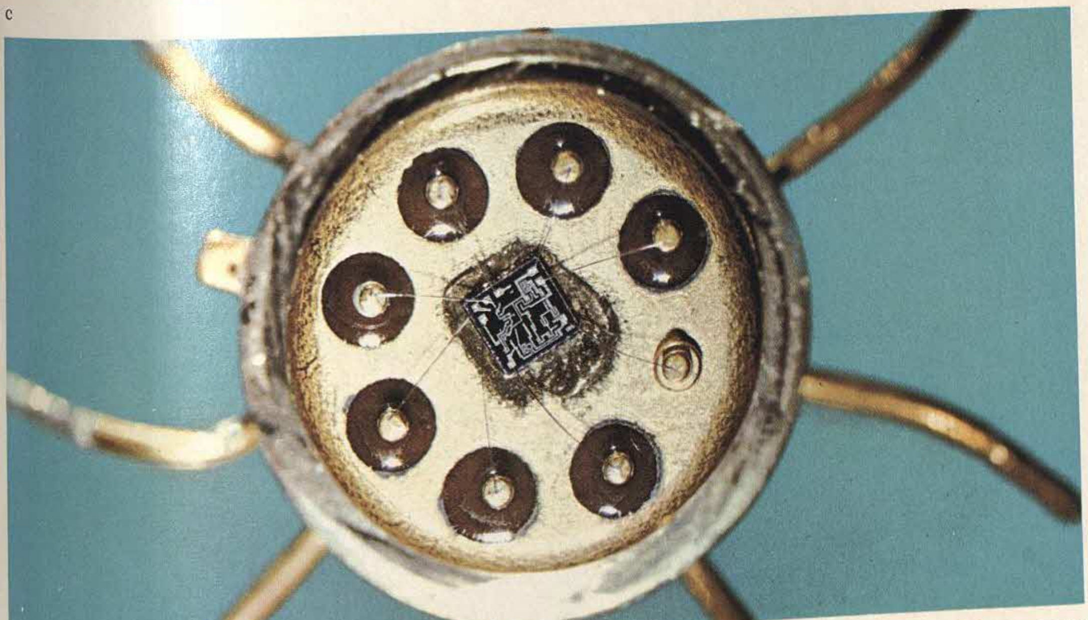
GEOMETRIC AND MECHANICAL CHARACTERISTICS

Transistors can be separated into two categories: power transistors, and those in which the output of power is not essential. In the former, a dissipator, or heat sink, is required around the semiconductive element to prevent excessive heating. The dissipating unit takes up an appreciable volume and these transistors are a third to half the size of corresponding vacuum tubes.

Nonpower transistors are used in electronic calculators where they serve to amplify electric signals of low voltage (generally between 3 and 6 V–volts). The currents involved are of the order of 1 μa (milliampere) or less.

Because the power is very small, dissipation equipment is not needed. Thus the dimensions are very small. Usually the active part has an area less than 1 mm^2 and a width less than 0.1 mm. These reduced dimensions require that the active element of the semiconductor be joined to a structure sufficiently large to be handled by a human operator. This

c



d

sistors capable of the same functions as the triode, each having a potential of 260 MHz. At the bottom is an integrated circuit of lesser output.

Illustration 1b shows open and closed views of an analog amplifier with an integrated circuit. Illustration 1c is the same analog amplifier enlarged 6.5 times. In Illustration 1d—on the left—is shown a tube for amplification up to 4 W and 500 MHz; to the right is a transistor capable of putting out up to 7 W at the same frequency. Sometimes the advantage of the

gives rise to a problem. Inside the transistor are thin wires connecting the semiconducting element with the fixed terminals on the protective covering. These wires are sensitive to vibration and are the only weakness of the transistor. Nevertheless, this sensitivity is less than that of any tube. The transistor is made of a leaf of extremely light semiconducting material fixed to a rigid support that is practically unbreakable. Vibrations usually break off the attachments at the terminals of the element without damaging the structure of the transistor itself.

In tubes, on the other hand, only a little deformation—caused by vibration—of the grid, plate, or filament is all that is needed to impair the functioning of the tube.

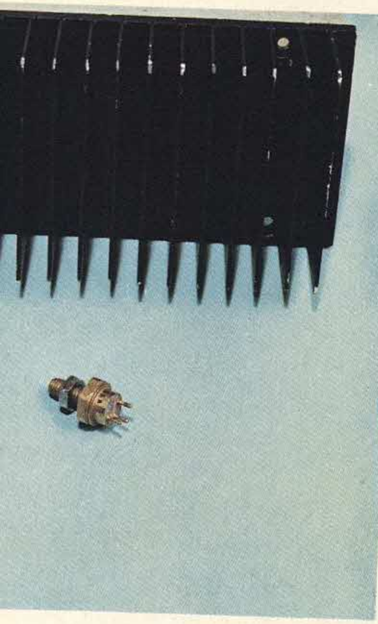
Mechanical solidity is much greater in plated micrological circuits; a transistor may have dimensions of a few thousandths of a millimeter. In these circuits, the connections to external terminals are formed by flat elements solidly fixed to the support of the circuit. They can withstand accelerations up to one hundred times the force of gravity and bursts of short duration up to 1,000 g's. An acceleration of 100 g's is equivalent to the head-on collision of two automobiles traveling at about 100 km/hr (about 62 mph).

The voltage used in plate microcircuits is still less—around 3 V; the currents are, in certain cases, of a few microamps. Under these conditions, the electrical energy dissipated is minimal and input provisions are simple, even in an apparatus that has a great many components. Moreover, these elements can withstand power overloads up to twice the normal voltage. That is, even if a circuit that has been designed for an input of 3 V is subjected to 6 V for a short time, the circuit is not damaged.

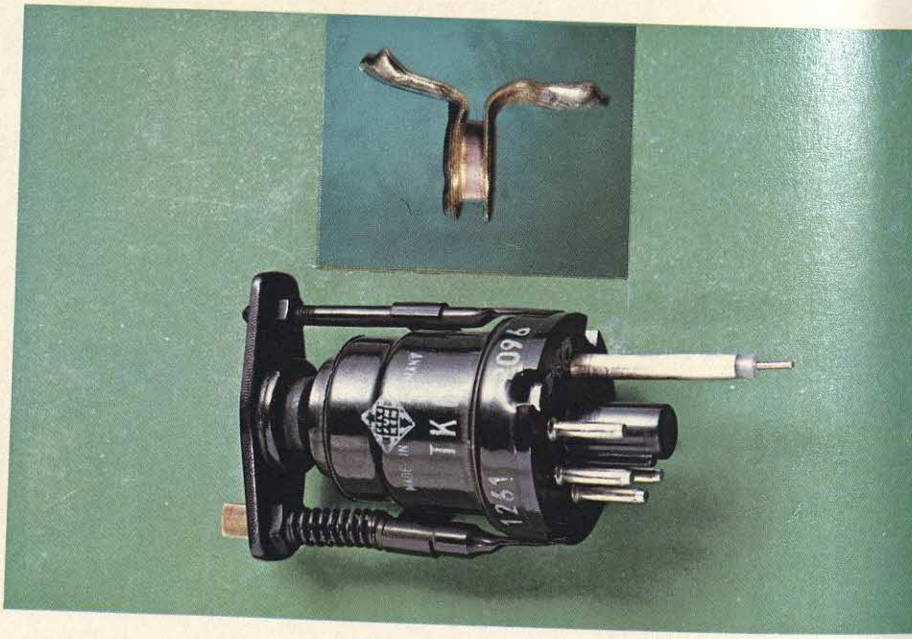
HOW TRANSISTORS FUNCTION

After the invention of the diode and the triode, the search began for ways to create an active circuit element whose functioning would not depend on thermionic emission. Initially, it was thought to exploit the Hall effect, a phenomenon in which the path of a current in a flat conductor is distorted by an electromotive force and by a properly applied magnetic field. Research in this direction was suspended when studies of the properties of semiconductors revealed possibilities of using semiconductors in the construction of new systems—primarily where transistors would substitute for vacuum tubes.

Transistors consist of a pair of junctions between semiconductors of two different types, P and N. P-type semiconducting material is positive, meaning that electrical conduction is due to a sparsity of electrons. N-type material is negative, meaning that conduction is by means of electrons. The junctions give rise to an opposing potential barrier. A voltage is applied to each of the two junctions, and a third voltage of the proper value must be applied to the central part to make the transistor function. The illustrations explain its function.



e



small-sized semiconductor components is nullified by the need to use a large heat dissipator to keep the temperature low, as shown in the top of Illustration 1d.

Illustration 1e shows a Klystron tube for 7,000 MHz. Above it is a diode tunnel, enlarged 6.5 times, which is used for the same function.

2

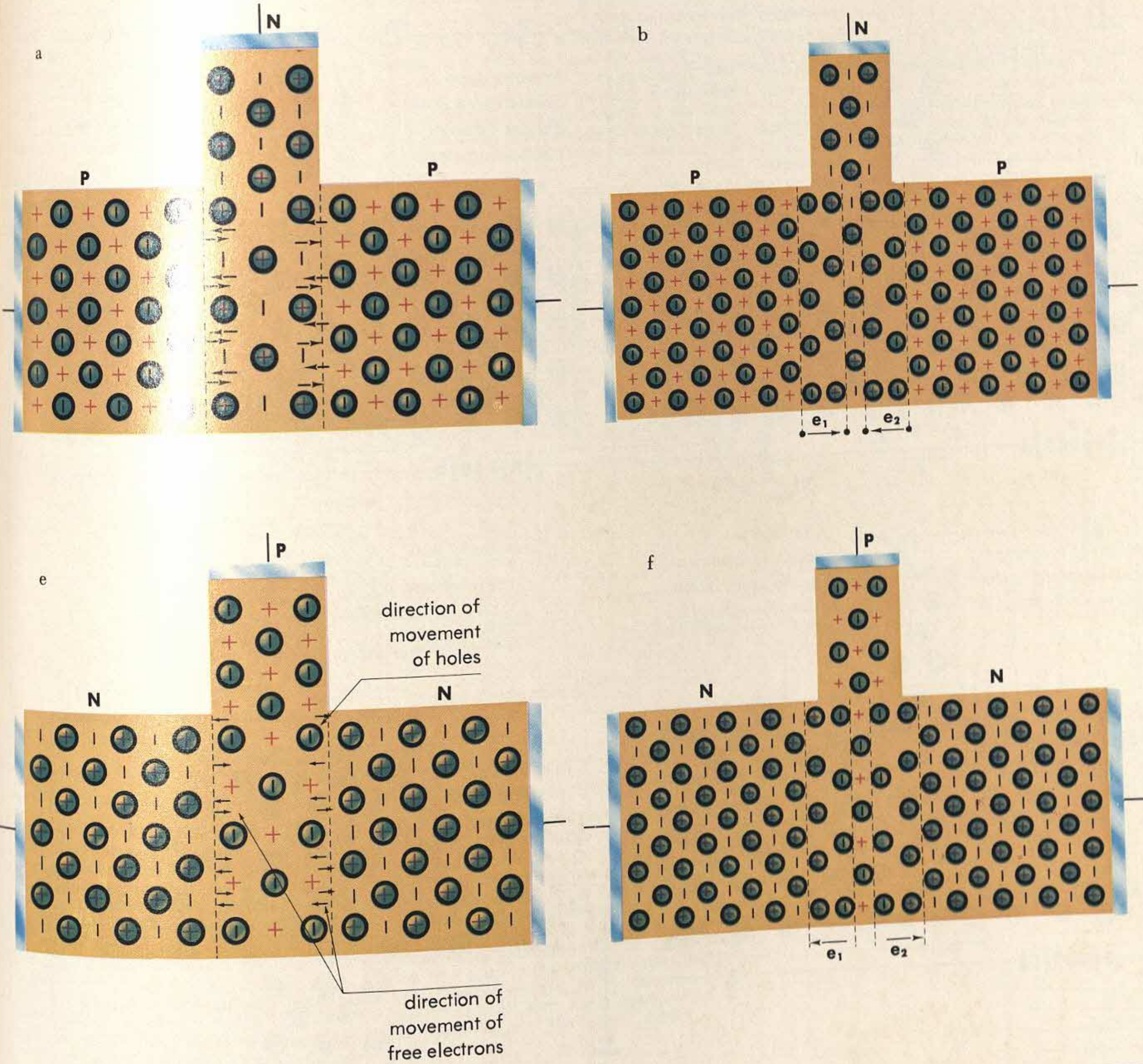
HOW TRANSISTORS FUNCTION—Illustrations 2a through 2d represent the electrical structure of a transistor of the P-N-P type.

The P regions are to the left and to the right of the N region, which is at the center. The three parts are formed into a single crystalline block. In the N region, the conductivity is due to electrons; in the P region the conductivity is due to holes; that is, a sparsity of electrons.

Illustration 2a shows the condition of charge at the moment the three layers of the block are joined and before the formation of potential barriers caused by the movement of the electrical charges across the borders of each layer (the movement is indicated by the pairs of arrows).

In Illustration 2b, electrons in the central N region spread into the lateral regions P, filling some of the holes. As a consequence of the contact of the regions two potential barriers are formed (similar to those in diodes): e_1 and e_2.

Illustration 2c shows how a transistor functions. When switch 1 is closed, the voltage applied to the two extremes P of the transistor reinforces the barrier e_2. When switch 2 is closed, a negative voltage is applied to the N region, which leads to the following: weakening of the e_1 barrier; movement of electrons from **B** to **A;** migration of holes from **A** to **B;**



movement of holes from region **B** to **C;** and a flow of current in the right-hand circuit.

In regions **B** and **C** of Illustration 2d the rupture of links between atomic pairs in the semiconductor gives rise to the moving charges or a current, accompanied by a thermal loss, which flows from emitter to collector. The current increases with temperature because the number of atomic linkages ruptured increases.

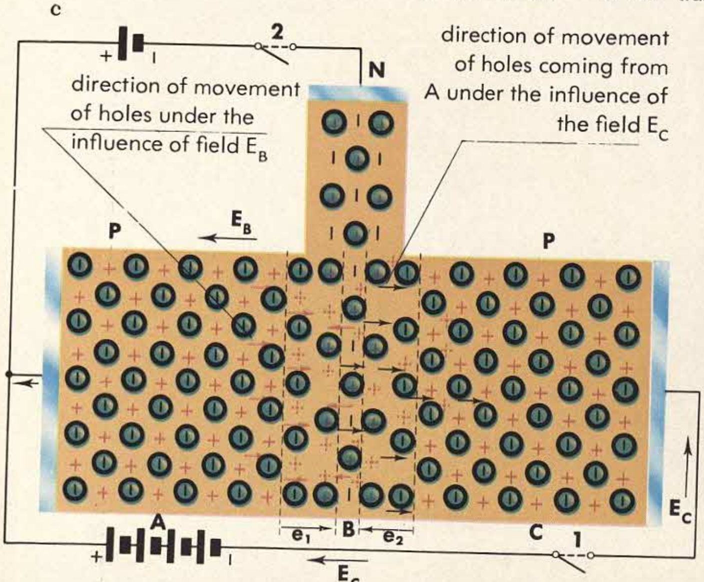
Illustrations 2e and 2f are analogous to 2a and 2b respectively, but show the condition of the semiconductors in a transistor of the NPN type. The nomenclature is the same.

Illustration 2g describes the function of the NPN transistor in a manner completely analogous to the PNP type (shown in Illustration 2c) except that the direction of the barriers, the input voltage of the external circuit, and the moving of electrical charges are reversed.

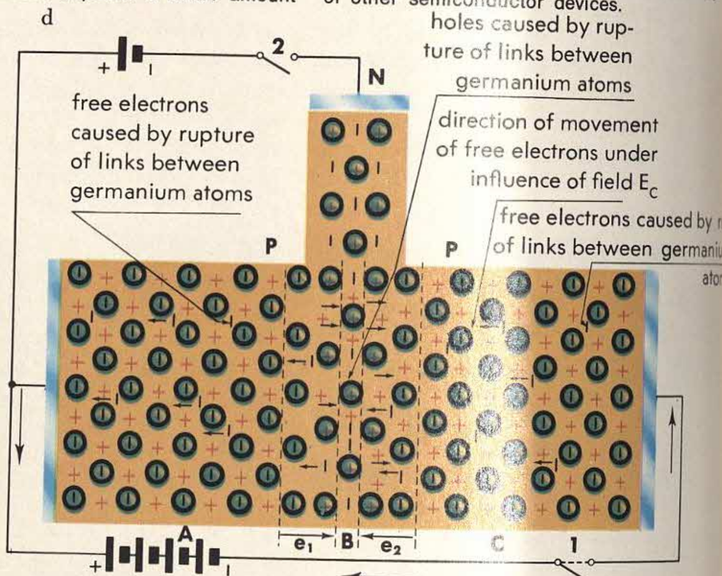
A PNP transistor deposits a small amount of the semiconductor P on a piece of semiconductor N about 50 μm thick. The junctions where the two types of semiconductors make contact have a thickness of about one μm. The lower part of Illustration 2h shows a modification of Illustration 2c, using similar nomenclature and arrows.

Transistor knowledge, techniques, and materials have been applied to make a variety of other semiconductor devices.

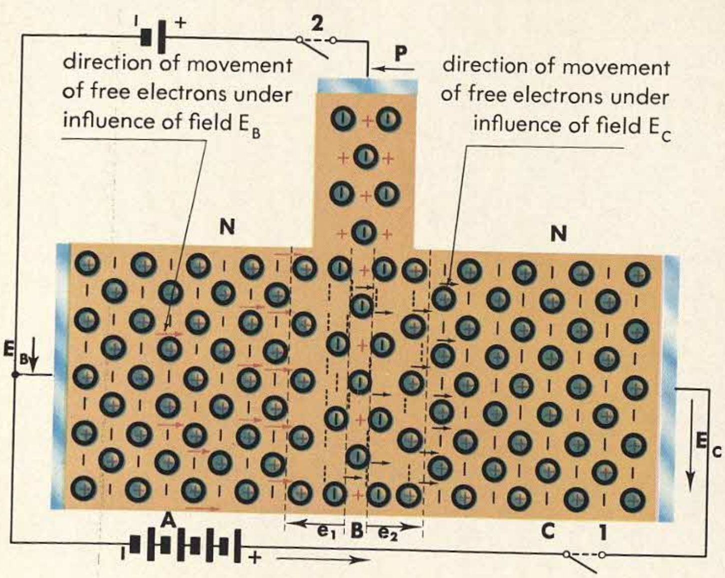
c



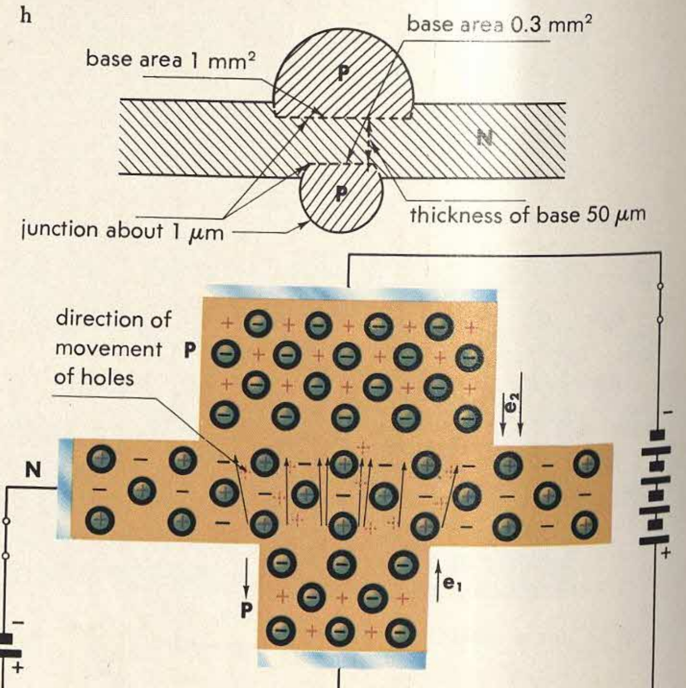
d



g



h



TRANSISTORS—II | design parameters

Several types of transistors are made, ranging from tiny, low-power units to devices designed to accommodate large currents at high voltages. Some transistors can work at low and intermediate frequencies, and others are suitable for very high frequencies.

A transistor itself is always small, but it is often enclosed in a casing to make it easy to handle. Various mechanical, electrical, chemical, and metallurgical techniques have been used to make transistors and other semiconductor devices.

The accompanying illustrations are concerned with transistor parameters and the circuits used to determine their characteristic curves. In actual practice, these curves are supplied by manufacturers so that the purchaser does not have to determine them before using the transistors in a particular circuit.

1

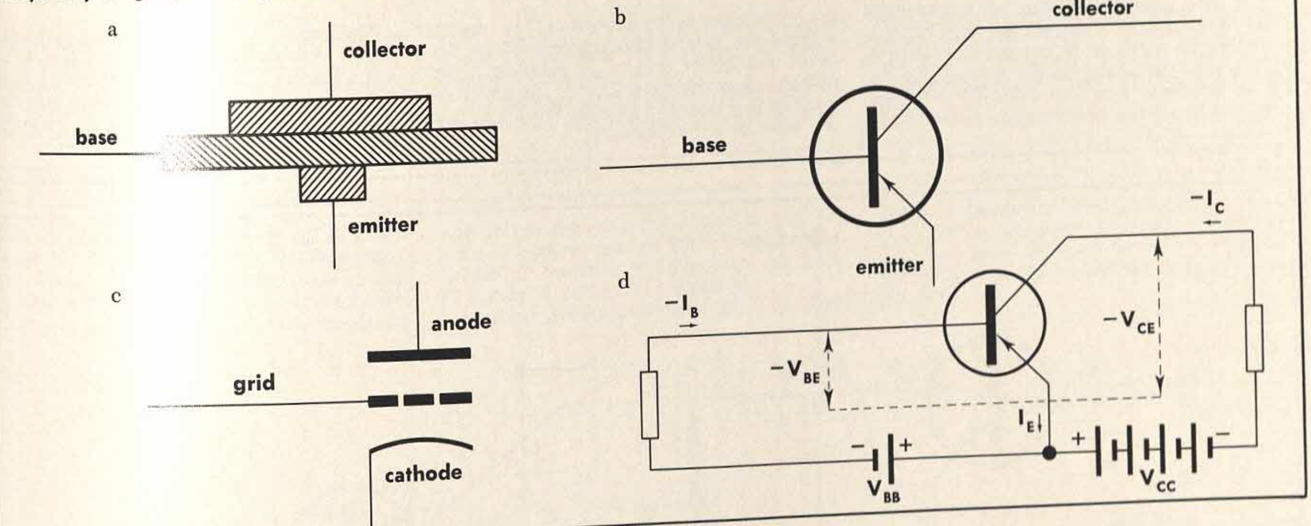
SOME CONVENTIONS—These diagrams show the symbols used to represent transistors, the terminology used for the various voltages and currents, and the conventions used for evaluating their magnitudes.

Illustration 1a represents a transistor schematically, in exactly the same way as it is constructed. The emitter and the collector are the two electrodes, and the current between them is controlled by the base.

Illustration 1b shows how transistors are normally indicated. (The circle is sometimes omitted.) Some variations of these symbols may be encountered, but their meaning is always easy to grasp. The symbol in Illustration 1b is the one for a PNP transistor; in the symbol for a NPN transistor, the arrow points away from the base, not toward it.

Illustration 1c shows the analogous elements of a transistor and a vacuum tube. The grid in the tube has the function of the base of a transistor. The cathode plays the part of the emitter; the anode takes the place of the collector.

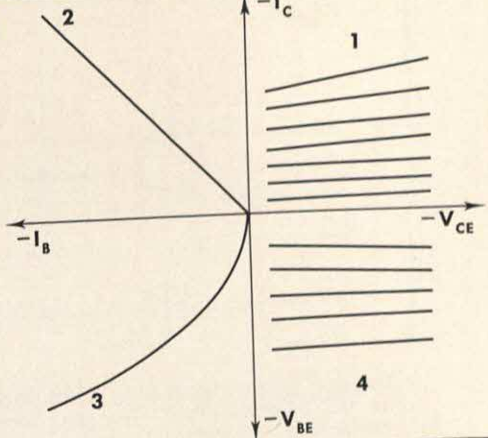
In Illustration 1d a transistor is shown in position in a circuit. The type of circuit shown is known as a common emitter circuit, because the emitter serves as a point of reference for establishing the voltages going to the base and collector. The meaning of the voltages shown in various parts of the circuit can be deduced from the symbols themselves. For example, $-V_{CE}$ indicates the collector-emitter voltage—the voltage difference between the collector and emitter, or $-(V_C - V_E)$. The two small rectangles included in the emitter-collector and emitter-base circuits represent loads. Currents are labeled by the same conventions as apply to voltages. If direct current (or current that varies only slowly) is used, capital letters are used for voltages and currents. If alternating current is used, lowercase letters are used for the instantaneous values of their magnitudes and capital letters are used to indicate their effective values.



CHARACTERISTIC CURVES OF A TRANSISTOR—This illustration shows the principal characteristic curves that define the behavior of a transistor and are used in the design of circuits. The four quadrants shown are really four separate graphs; the semiaxes have different scales both in units of measurement and quantities measured ($-I_C$, $-I_B$, $-V_{CE}$, and $-V_{BE}$).

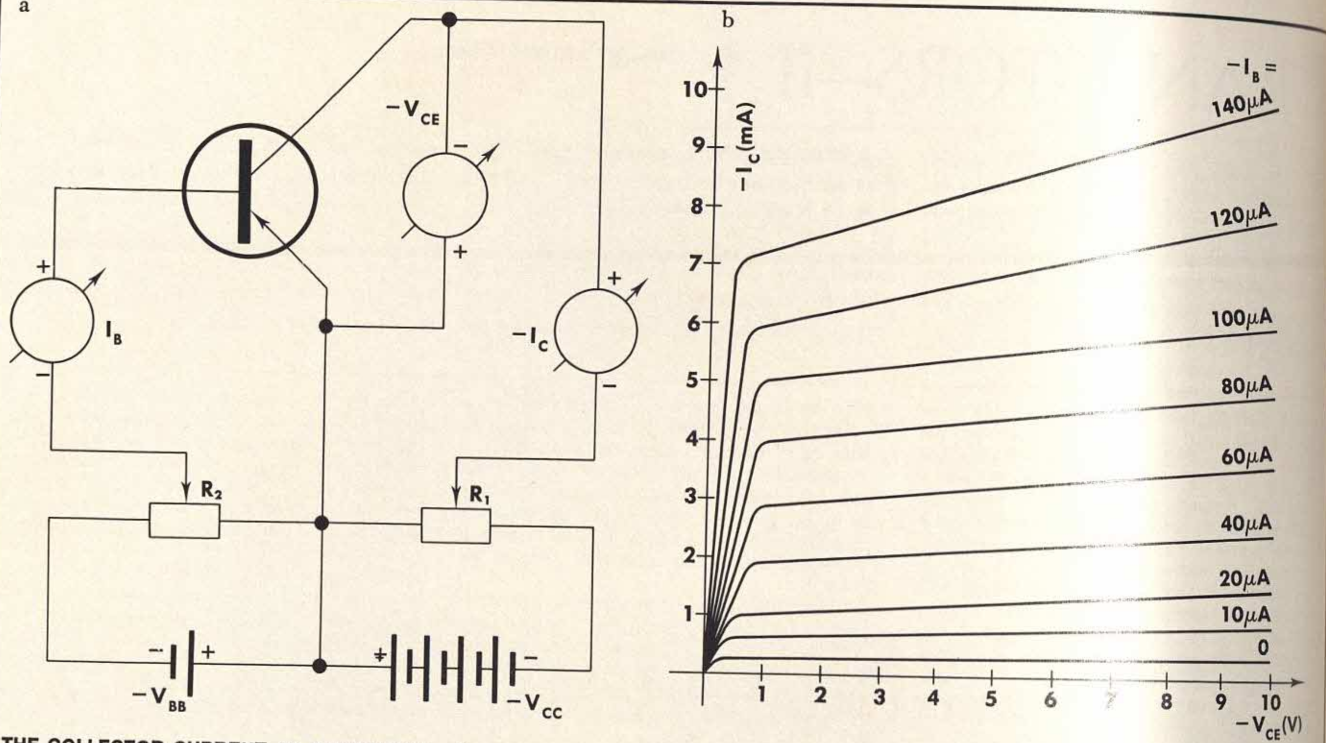
In the upper right-hand quadrant the curves represent the variation of collector current in terms of collector-emitter voltage. Each curve corresponds to a particular value of the base current. The upper left-hand quadrant shows how the collector current varies with respect to the base current. In the lower left-hand quadrant, the base current is plotted against the base-emitter voltage. It can be seen that this variation is not linear like the first or second quadrants. In the lower right-hand quadrant are the characteristic curves showing how base-emitter voltage varies with collector-emitter voltage. As in the upper right-hand quadrant, each curve corresponds to a particular value of the base current.

2



3

a b

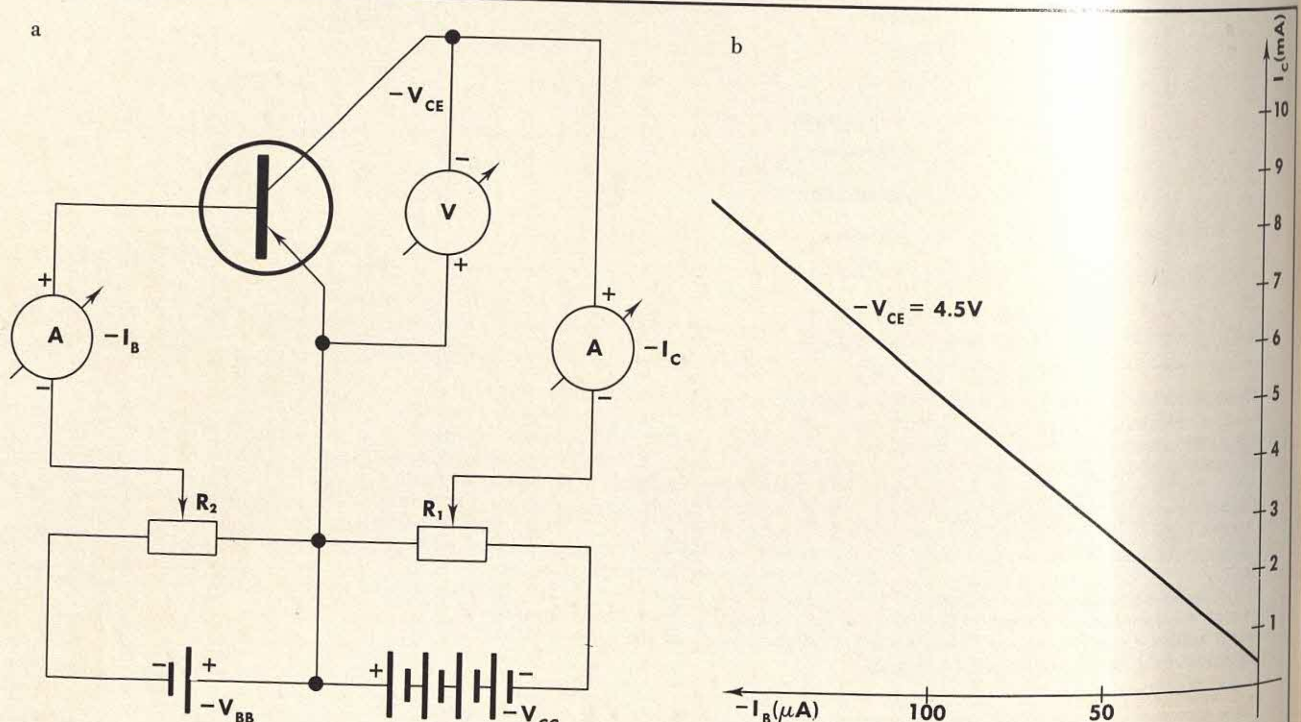


THE COLLECTOR CURRENT AS A FUNCTION OF THE COLLECTOR-EMITTER VOLTAGE—The circuit shown in Illustration 3a is used to determine the curves shown in Illustration 3b. Two ammeters and one voltmeter are used. R_1 and R_2 are two potentiometers used to vary the parameters being calibrated. The voltage sources are symbolically represented as batteries, but better results are obtained if stabilized supply devices are used. The ammeters must be very sensitive because current I_B is of the order of a microampere.

The graph of collector currents as a function of collector-emitter voltage is seen to be a straight line after the short but steeply rising initial stage is passed. The slope of the characteristic curves increases with increasing base current.

4

a b

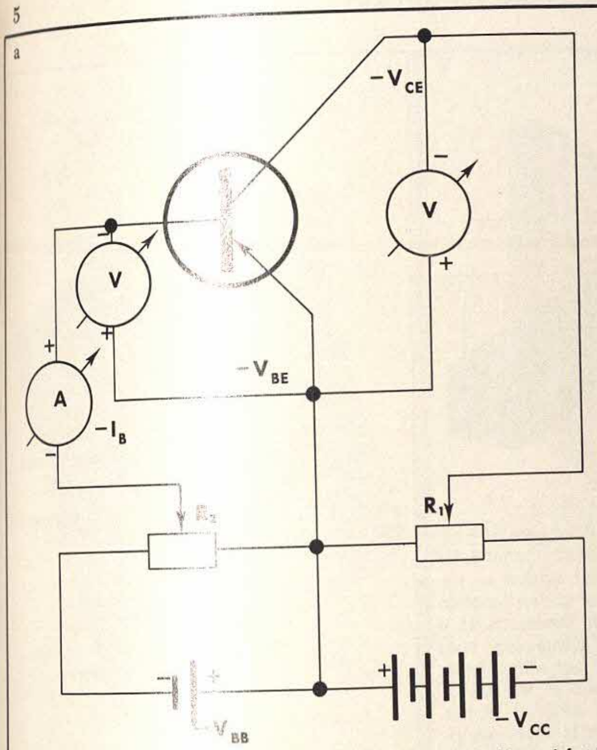


THE COLLECTOR CURRENT AS A FUNCTION OF THE BASE CURRENT—The same circuit as was used in Illustration 3 is shown here on the left as Illustration 4a. Some transistors require the use of special electronic voltmeters with a very high input impedance that will not affect the value of the voltage to be measured.

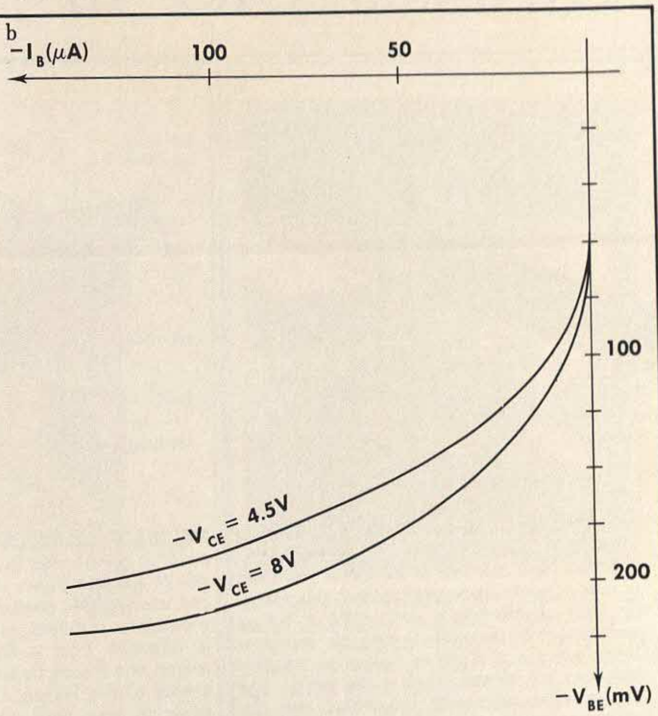
In this illustration, potentiometer R_2 varies the base current and the voltage difference between collector and emitter is kept constant. The characteristic curve (Illustration 4b) is a straight line in this case. If the emitter-collector voltage is varied, the corresponding characteristic curves vary only slightly from each other because these curves are not greatly affected by the emitter-collector voltage.

5

a



b



THE BASE CURRENT AS A FUNCTION OF THE BASE-EMITTER VOLTAGE—The voltmeter used to measure the base-emitter voltage is an electronic device because the voltages to be measured are small (only fractions of a volt). Except for this, the circuit shown in Illustration 5a resembles the one shown in the preceding illustrations.

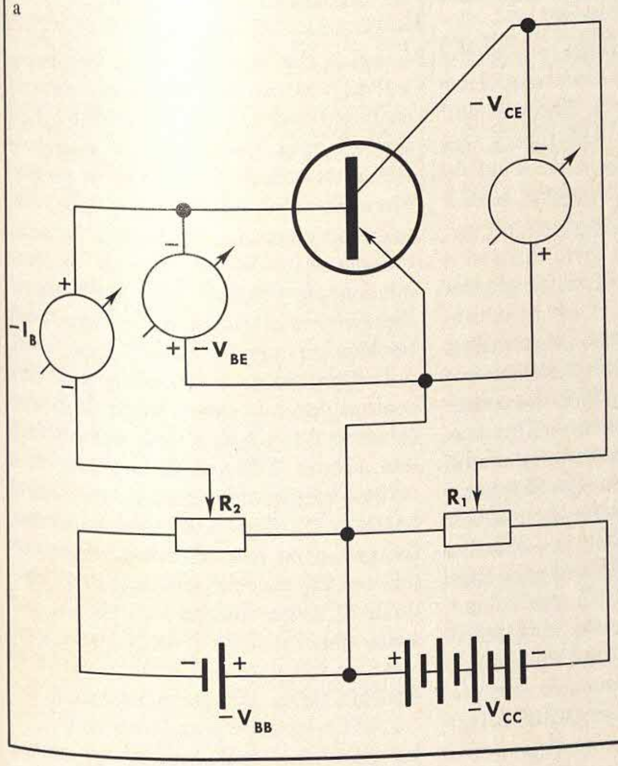
The different curves shown in Illustration 5b are obtained as a result of different values of the collector-emitter voltage. Potentiometer R_1 is used to set this voltage at the different values. The corresponding characteristic curves are then determined; they are very close to a parabola.

6

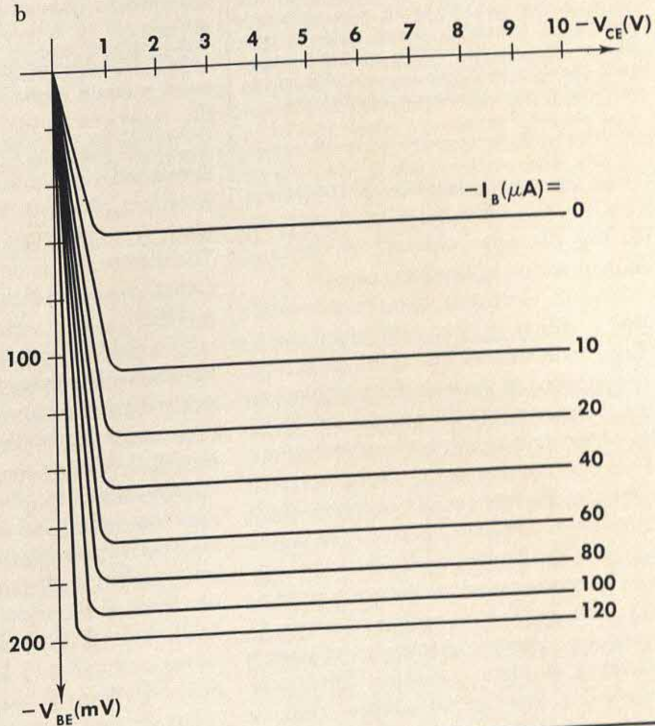
THE BASE-EMITTER VOLTAGE AS A FUNCTION OF THE COLLECTOR-EMITTER VOLTAGE—The same circuit is used here as was shown in Illustration 5a. The characteristic curves shown in Illustration 6b are strongly influenced by the value of the base current. This value, therefore, must be fixed for each curve. Except for the short, initial portion this family of curves is practically horizontal.

Knowing the characteristic of a transistor makes it possible to predict its behavior in a circuit. However, other parameters—such as the impedance between terminals—must be determined to size the various components of the circuit.

a

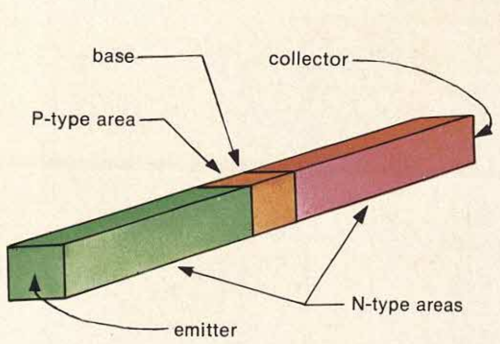


b



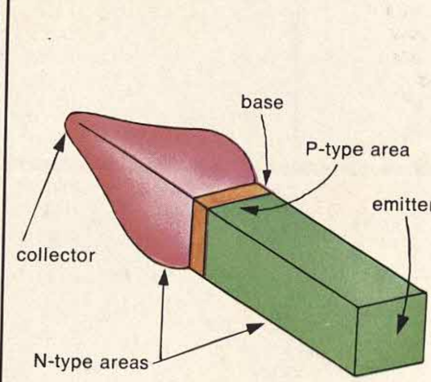
TRANSISTORS—III | modern structure and production techniques

1



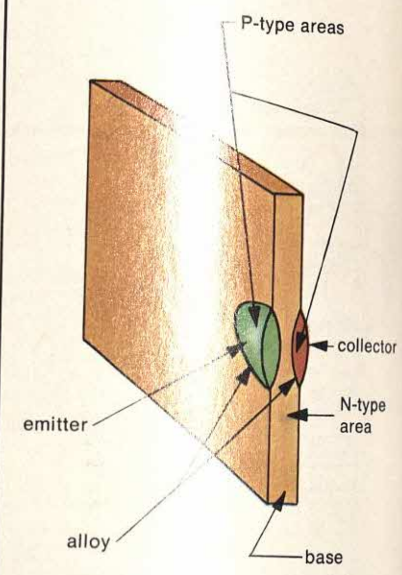
POINT CONTACT TRANSISTORS AND GROWN JUNCTION TRANSISTORS—The point contact transistor was the first to be constructed but is now obsolete. It consists of two thin metallic wires attached to a small cube of N-type semiconductor material. During the manufacturing process, the unit is treated so that P-N junctions are formed close to the points. One of the wires represents the emitter and the other the collector, while the small cube of semiconductor material is the base of the transistor. The transistor shown here is a grown junction transistor obtained by causing the P-N junctions to grow during the original process of forming the semiconductor monocrystal. It consists of a parallelepiped of germanium or silicon in which the two P-N junctions bound a thin base zone. The two ends of the parallelepiped constitute the emitter and the collector. The production techniques presently in use are based on the successive addition of appropriate impurities to the fused material from which the monocrystal is being grown, thus resulting in two N-type layers separated by an interposed P-type layer. The crystal is then cut into small bars, each representing a transistor. These types of transistors, although still widely used, have been made obsolete by others that are obtained by means of more sophisticated techniques.

2



MELTBACK TRANSISTORS—This type of transistor is constructed from a small, rectangular bar of semiconductor material similar to the one used for the production of grown junction transistors. It differs, however, inasmuch as it is obtained from a crystal containing both N-type and P-type impurities, but with a prevalence of the N-type. One end of the bar is made to fuse until a small drop of molten material is formed; this drop is then recrystallized. During the solidification the P-type elements become concentrated at the boundary between the fused and the unfused material and thus form a thin P-type base layer. The characteristics of this transistor are similar to those of the grown junction type but exhibit a distinct improvement in behavior at high frequencies.

3



ALLOY JUNCTION TRANSISTORS—These are now used extensively and are generally germanium transistors of the P-N-P type. The base is a small plate of N-type material. Small disks of aluminum or indium are fixed to the upper and lower faces of this plate and then heated until they form an alloy with the germanium of the base. The transition zones constitute the emitter and collector junctions. The same method can be used to obtain N-P-N transistors; in that case the alloying materials are of lead or indium with additions of arsenic. The thinness of the base limits the working voltage but at the same time improves the behavior of the transistor at high frequencies.

The transistor was invented in 1948. Since that time, its manufacturers have been looking for new methods to improve characteristics and reduce costs.

Modern electronic computers contain half a million or more transistors, particularly transistors of the silicon plate type. The creation of such complex equipment is a direct result of the enormous reliability of the most recent types of transistors. These permit computers to be operated continuously, without excessively frequent breakdowns. Present efforts in the search for new production techniques are primarily concentrated on lowering production costs and increasing the reliability of transistors during operation. Operating lifetimes of plate transistors have now reached a limit of 20 million hours, a value unthinkable until a few years ago. To obtain proper appreciation of the exceptional capabilities now achieved in transistor technology, it is necessary to take a look at the various stages that have led to present processes and techniques.

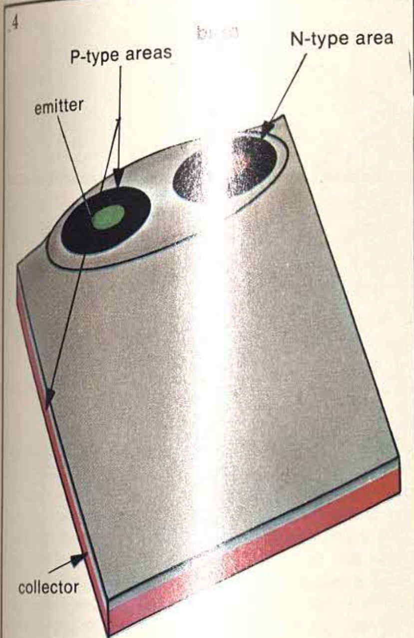
The first transistor, which was of the point contact type, was constructed by the American physicists J. Bardeen and W. H. Brattain in 1948. The theory was developed simultaneously by the English-American physicist, W. Shockley, particularly for transistors of the junction type. The discoveries made by these three men gained them the Nobel Prize for physics in 1956.

The field of application of the first transistors was very limited because it was not possible to reproduce the conditions indicated by the theory. This was essentially due to the inadequacy of the manufacturing techniques. The first junction transistors, for example, attained a maximum operating frequency of 20 MHz (megahertz) and were limited to a base thickness of the order of 0.1 mm (about 0.004 in.). Base thicknesses now reach values as small as 1 to 2 μm (microns), thus allowing the transistors to operate at frequencies of the order of 1,000 MHz. One of the factors that has made it possible to achieve this increased efficiency is found in the direct control (morphological as well as electrical) now directly exercised over the junctions. Both optical and electron microscopes (particularly the latter) have been used to do this. More recently, it has been done with scanning electron microscopes. In order to appreciate the importance of these microscopic controls, the dimensions of the active material in a transistor must be kept in mind. The old types, with grown junctions, made use of small silicon bars, each having a length of 2.5 mm (about 0.1 in.) and a section of 0.5×0.5 mm (about 0.02×0.02 in.), that is, a volume of the order of 0.6 mm³ (about 0.0004 in.³). The most recent diffusion transistors, on the other hand, have volumes of the order of 0.01 mm³—the same order of magnitude as, say, the head of a tiny pin.

FORMATION OF JUNCTIONS

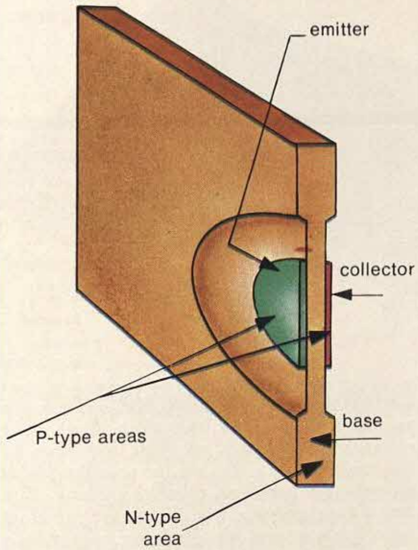
Junctions may be obtained by adding impurities (doping substances) during

4



POST-ALLOY DIFFUSION TRANSISTORS—This transistor, generally known by the initials PADT, is constructed on a small lamina of P-type semiconductor (usually of germanium). A prediffusion process creates a certain depth of N-type material on the surface of the lamina. Two thin, metallic disks are then applied to the N-side of the lamina: the disk that will become the base contains N-type impurities; the one that will become the emitter contains both P-type and N-type impurities. The lamina itself eventually acts as the collector. The whole is then heated under controlled conditions. The germanium softens and the base and emitter impurities diffuse into it. N-type impurities are chosen for their great speed of diffusion and form an N-type layer as they penetrate the lamina. The P-type impurities of the emitter diffuse more slowly. After cooling and recrystallization, the emitter region consists mostly of P-type material separated from the collector (which also consists of P-type material) by a thin layer of diffused, N-type material that constitutes the base.

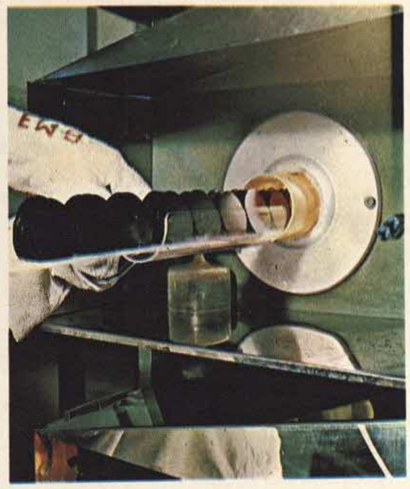
5



SURFACE BARRIER TRANSISTORS—The production process for this transistor is of the electrochemical type and permits the formation of an extremely thin base region. A small plate of N-type semiconductor material is placed between two extremely fine jets of a metallic electrolyte solution. Under action of a direct voltage, the semiconductor material is notched by the solution. When the desired depth is obtained, the polarity of the direct current is inverted and the solution deposits small spots of metal on the sides opposite to the ones it attacked previously. These metallic spots become the emitter and collector electrodes, while the attacked plate becomes the base. The transistor obtained in this manner can then be heated in an oven; this produces a diffusion of the deposited metal and results in a surface barrier transistor of the diffused type (Illustration 6).

All surface barrier transistors are characterized by excellent high-frequency response, but this is offset by their having relatively low working voltages.

6



DIFFUSED TRANSISTORS—Diffusion has made possible considerable improvement of transistor characteristics. This technique is based on the use of silicon as starting material. Silicon has a number of advantages when compared with germanium, including a greater maximum working temperature and a smaller dispersion current. When the silicon is heated to a high temperature and exposed to a source of impurity, the atoms of the impurity diffuse into the semiconductor material (somewhat as dye diffuses in water). A P-N junction of the diffused type, for example, can be obtained by diffusing N-type impurities such as phosphorus or antimony in a layer of P-type silicon. If the temperature and the diffusion time are accurately controlled it is possible to obtain a diffusion of exactly the desired depth. Consequently, the junction will have the required thickness. Moreover, a layer of oxide can be created on the surface of the silicon bar. This oxide layer has the property of permitting the atoms of some materials to diffuse into the underlying mass while other materials are stopped, or masked. The masking effect also offers a solution to the problem of obtaining very small junctions of a complex form. In practice, this is achieved by means of photoengraving techniques. The oxide layer is removed from the chosen areas of the semiconductor, and the impurities then diffuse only into these particular areas because the oxide mask covering the remainder prevents it from being penetrated.

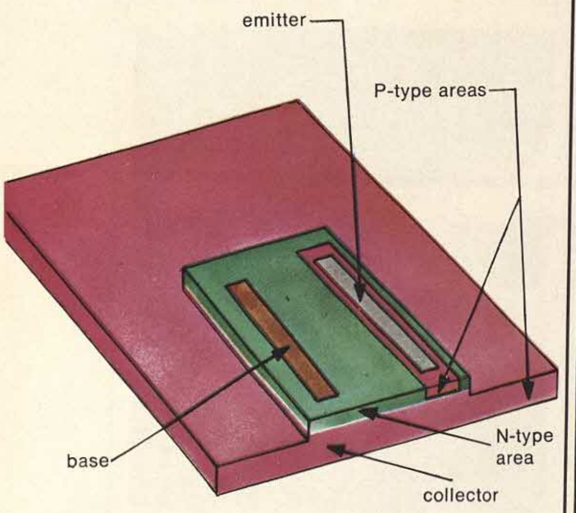
The photograph shows a silicon diffusion oven.

the growth process that is employed for obtaining a monocrystal of the semiconductor from the fused material. When producing junction transistors with this technique, fused material with a low concentration of N-type impurities is used. At a certain point during the growth process, a certain quantity of P-type material is added to the fused material to ensure that the part of the crystal then being formed will be of the P-type. After a relatively short time, a strong dose of N-type material is then added to the fused material. This ensures that the P-type part of the crystal will be thin (transistor base) and that the remaining part of the crystal will be of the N-type and have a low resistivity. This latter part of the crystal constitutes the emitter of the transistor. The part of the crystal that is formed first has a relatively high resistivity because of the small concentration of N-type material, and it constitutes the collector of the transistor. When this technique is employed, the density of the impurities varies gradually along the length of the crystal. This technique is preferred for N-P-N transistors made of germanium or silicon.

FORMATION OF JUNCTIONS BY ALLOYAGE

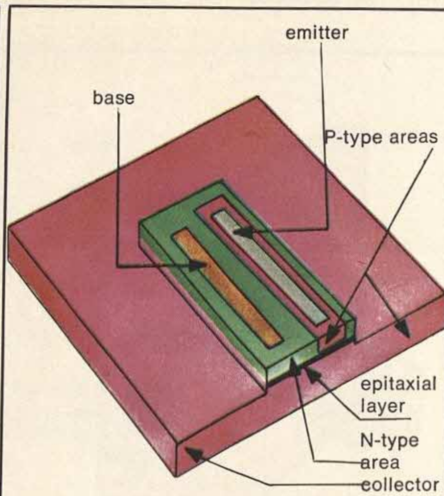
This method, now used widely in industry, employs indium or indium-gallium. A small sphere of indium having a diameter of a few tenths of a millimeter is placed on a thin germanium lamina with a rather high N-type resistivity, and the whole is then heated to a temperature of about 500° C (about 932° F). At this temperature a small drop forms on the germanium and sinks into it—the depth of the depression formed depends on the quantity of indium used. When the temperature is subsequently diminished, a part of the germanium recrystallizes as P-type germanium with a low resistivity. A very clearly defined junction is obtained—this is the emitter junction of the

7



MESA TRANSISTORS — This name derives from the appearance of these transistors when examined under a microscope. They resemble the flat-topped hills known as mesas (a Spanish word) characteristic of the Southwestern part of the United States. Several processes are used in the production of mesa transistors. In one process, a layer of semiconductor material serves as a collector. A thin film of N-type impurity is then diffused in vapor form on the P-type material; the base is formed in this way. The P-type emitter region is obtained by means of an evaporation process under vacuum, or by means of an alloying process. The construction of the transistor is then completed by allowing a chemical attack to take place all around the base-emitter region, thereby obtaining the characteristic form of a mesa. These transistors can also be produced by means of double diffusion processes, in which the silicon is exposed to two different diffusions, each of which is preceded by appropriate masking and engraving. The first diffusion produces the base-collector junction, which is rather large; the second diffusion produces the emitter-base junction, which is smaller than the other junction and is placed very accurately above it. This technique has made it possible to equal (and even exceed) the already excellent high-frequency characteristics of germanium transistors.

8



EPITAXIAL MESA TRANSISTORS—These are very similar to mesa transistors when seen under the microscope, but differ from them by the formation of a thin film (or epitaxial layer) between the base and collector regions. This film is homogeneous with that of the principal collector body and represents an intermediate collector electrode. By means of these films it is possible to reduce the voltage drop at saturation and to improve power performance.

9

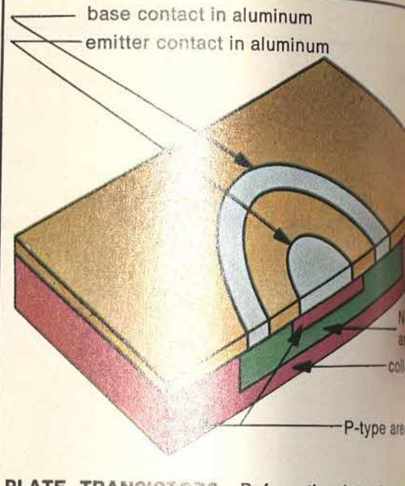


PLATE TRANSISTORS—Before the introduction of the plating process, the transistors produced by other techniques had their junctions exposed to the surrounding environment. The immediate consequence of this exposure was that the surface junctions had characteristics and properties that differed from similar junctions embodied in the mass of the semiconductor. This caused considerable instability and called not only for the most rigorous and extreme cleanliness during manufacture, but also required each transistor to be hermetically enclosed. The plating process has eliminated all the problems connected with the exposure of the junctions. The process is begun by first forming, at an extremely high temperature, a layer of silicon dioxide on the plates of the material. The semiconductor is thus sealed within the layer of silica. The diffusions are then carried out in such a way that the junctions form below the layer of oxide and never come into contact with the surface.

The experimental characteristics of these transistors are extraordinarily close to the theoretical ones. A layer of N-type material (silicon, for example) serves as collector. A film of oxide is formed on its surface and this serves as protection from diffusion of impurities in the material. The base and emitter regions are constructed by means of diffusion of P-type and N-type impurities in successive stages, after appropriate removal of the oxide film. Subsequently, metallic aluminum is deposited on both the base and the emitter regions, thereby lowering the resistance of the contacts. Lastly, the junctions are covered by a final layer of oxide that prevents contamination and increases electrical stability. The manufacture of plate transistors is divided into the following stages: (1) preparation of the silicon—growing the crystal and cutting and lapping the slices; (2) oxidation—preparation of the slices for diffusion (the processing is carried out simultaneously on many slices from which thousands of transistors are obtained); (3) masking and diffusion—removal of the oxide from the areas where the diffusion is to take place, formation of the base region, and masking and formation of the emitter regions; (4) metallization and formation of the contact areas; and (5) assembly and check—with subsequent separation of the individual transistors, welding to the base, connection of the contact areas to the external wires with extremely thin gold wires, and soldering of the cap. Naturally, electrical and mechanical tests are carried out on the finished transistors.

transistor. At the other end of the bar of germanium the collector junction is produced in an analogous manner. In this way the resistivities of the emitter and the collector of the transistor are almost identical. This technique is generally preferred for the production of P-N-P germanium transistors.

FORMATION OF JUNCTIONS BY DIFFUSION

Junctions may be obtained by exploiting the diffusion of impurities that occurs when a semiconductor crystal (with small quantities of doping substances on the surface) is heated to a high temperature. This technique is widely used for obtaining diodes and for applying the different diffusion speeds of P-type and N-type impurities—for obtaining a type of transistor known as a diffused junction transistor.

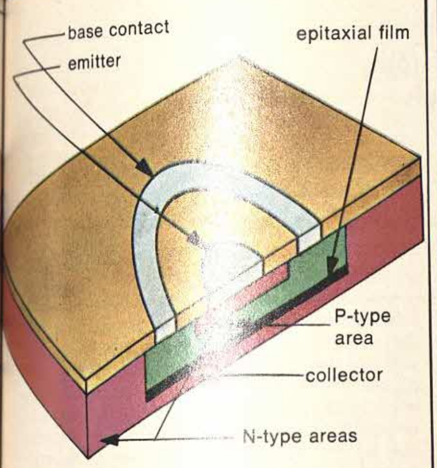
These general principles are applied in order to obtain the various types of transistors shown in the illustrations.

The plating process is the most important of the various processes illustrated because it makes possible the construction of integrated circuits. The development of integrated circuits represents an enormous leap toward microminiaturization, greater reliability in operation, lowering of absorbed power, and reduction of production costs. The most advanced techniques are:

1. Subminiature mounting of conventional components: the dimensions of each electronic component are reduced to the maximum extent possible. These microcomponents are then mounted in microcircuits by means of automated processes. There is now a tendency to abandon this method because of its high cost.

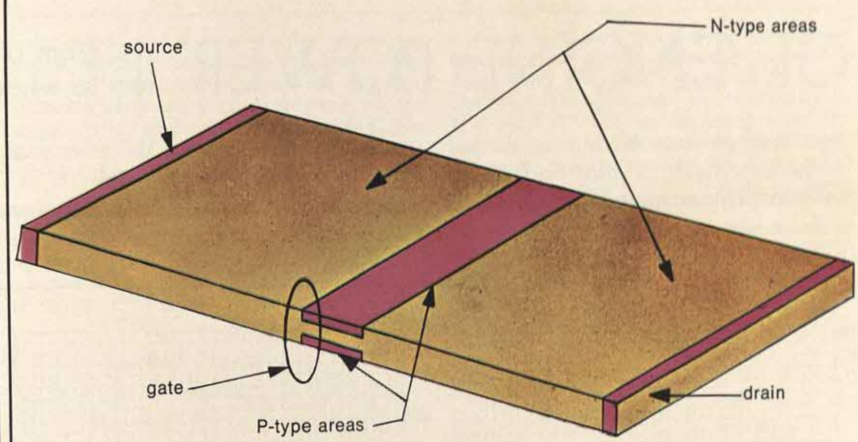
2. Thin-film circuits obtained by means of evaporation: the components, resistances and condensers, for example, are obtained by including thin layers of various materials. The attempts to produce

10



EPITAXIAL PLATE TRANSISTORS — When applied to plate transistors, the epitaxial process consists of first depositing a further layer of extremely pure material on the silicon slices. This is achieved by decomposition of a gaseous compound of silicon. The masking and subsequent diffusion processes are then carried out on this epitaxial stratum.

11



FIELD EFFECT TRANSISTORS—The construction and the operating principles of this transistor are extremely unique. Even its electrodes have different names; they are generally called source, gate, and drain, rather than emitter, base and collector. The input impedance is extremely high and, in a certain sense, the transistor behaves like a low-voltage vacuum tube. A typical field effect transistor (FET) consists of a small bar of N-type semiconductor material with P-type impurities applied to its opposing sides. This creates two P-N junctions with an N-type channel between the two P-type regions. Metallic contacts are constructed at the opposite ends of the bar. These serve as the source and drain electrodes; the two P-type regions constitute the control electrode, or gate.

active elements like transistors have not been too successful.

3. Integrated semiconductor circuits: these have been introduced by means of an extension of the plating technique and are proving to be more suitable for covering present and future needs. All the circuit functions of transistors—those of diodes, condensers, and resistances—are obtained by means of appropriate combinations of P-N junctions and semiconductor zones of different resistivity.

4. Functional blocks: the circuit functions are carried out by homogeneous elements of the material and the complex functions are obtained by the union of simple elements. The identity of the individual electronic components is lost. A functional combination of inductance, capacity, and resistance, for example, is performed by a quartz crystal. The development of this sector is closely connected with the discovery of new materials or new properties, but, in any case, it appears as if these devices will be unable to carry out all the functions that are required in electronics.

CONTACTS

Another problem of considerable importance has to be solved in the construction of the contacts. The external contacts are generally made of tinned copper, but the internal ones must be constructed so as to avoid the contact rectification phenomena that would occur at a metal-semiconductor junction. This explains the great variety of metals and alloys used for making the internal contacts of transistors.

Eliminating all traces of rectifying action at the contact is not always simple, but it may be accomplished in a number of ways.

Alloying.—This process is similar to the making of an alloy P-N junction except that the type of impurity used is the same as the substrate rather than opposite; thus, instead of a rectifying junction, a low resistance contact results.

Diffusion.—Just as in alloying when the diffusate is of the same rather than the opposite conductivity type, a good ohmic contact results. After diffusion an evaporated metal contact is alloyed in.

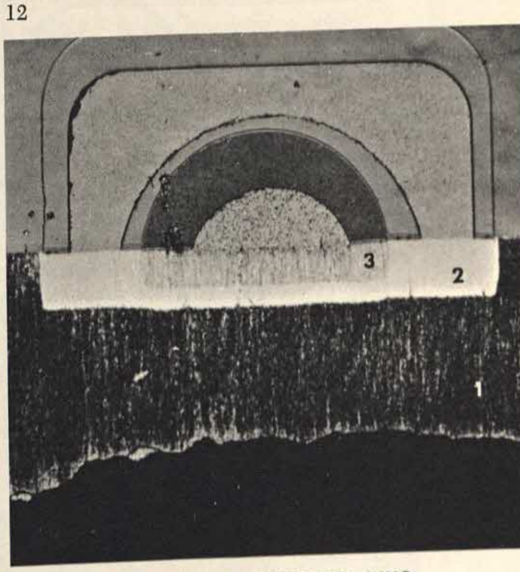
Plating onto an Abraded Surface.—Here the mechanical working of the material so reduces the minority carrier lifetime that rectification is destroyed.

Welding, Bonding, or Soldering.—Since these methods involve formation of a liquid phase during the process, they are closely related to alloying.

Point Contact to an Abraded Surface.—This method is often unsatisfactory because the resistance is too high. However, if the current is kept small, a point contact may be used for test purposes.

Epitaxial Layers. — A controlled thin layer of semiconductor suitable for making high-speed transistors may be made by deposition from a hot gas onto a suitable substrate. Transistors made on epitaxial material usually have superior properties because of their lower internal resistance.

12

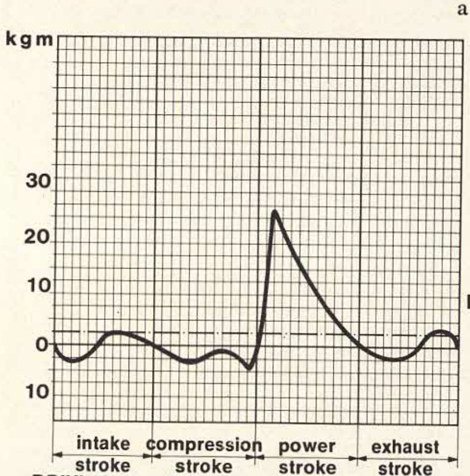


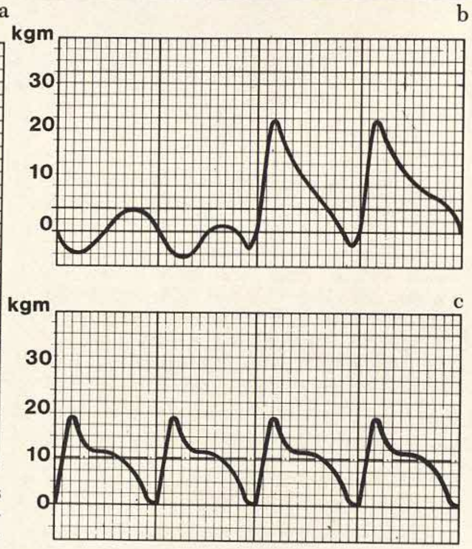
SECTION THROUGH A TRANSISTOR JUNCTION—The junctions in this photograph were made visible by lapping the transistor at an angle of 5°. The photograph was taken while the transistor was strongly illuminated by white light, with an exposure time of 30 seconds. The collector **1**, the base **2**, and the emitter **3** can be seen clearly. This is an N-P-N type diffusion brought about by treatment with a solution of hydrated copper sulfate ($CuSO_4 \cdot 5H_2O$) and concentrated hydrofluoric acid.

THE TRANSMISSION OF ENGINE POWER | from piston to wheels

Another article describes how an internal combustion engine is made and how it functions. This article takes up a related problem—the transmission of force developed by the engine to the wheels of the vehicle. In essence, transmission involves the conversion of the rectilinear motion of the piston into the rotational motion of the wheels. On the surface, this seems to be rather simple; yet, just as soon as the details of the problem are examined, it no longer appears so simple.

1



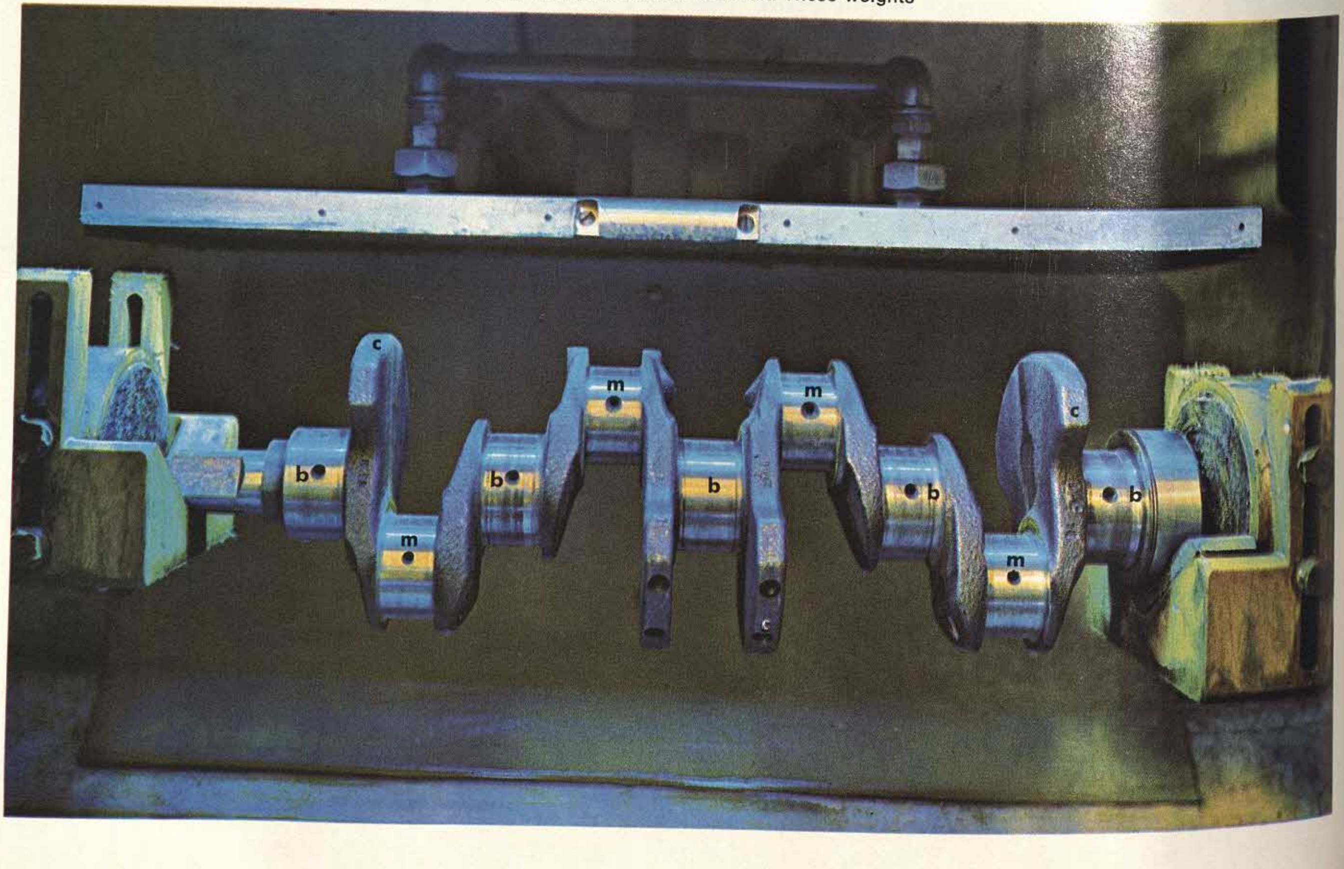


DRIVING TORQUE—During the course of its travel the piston is subjected to a force directed along the axis. This force has two components: one component acts along the axis of the connecting rod; the other acts perpendicular to the cylinder walls. This latter component tends to push the piston against the walls of the cylinder. It is, therefore, a source of power loss as a result of friction between the piston and the cylinder wall. The component acting along the axis of the connecting rod, on the other hand, is a driving force. It exerts a torque on the engine shaft—the driving torque. This driving torque has a negative value during the three idle strokes of the piston; its value is positive during the power stroke. The variations of the driving torque in a single-cylinder engine are shown in Graph 1a; those in a two-cylinder engine, in Graph 1b; and those in a four-cylinder engine, in Graph 1c.

As the graphs show, an increase in the number of cylinders reduces the negative portion of the driving torque. The dotted horizontal line in each graph represents average driving torque. The engine shaft transmits the driving torque to the gearbox and then to the wheels.

2

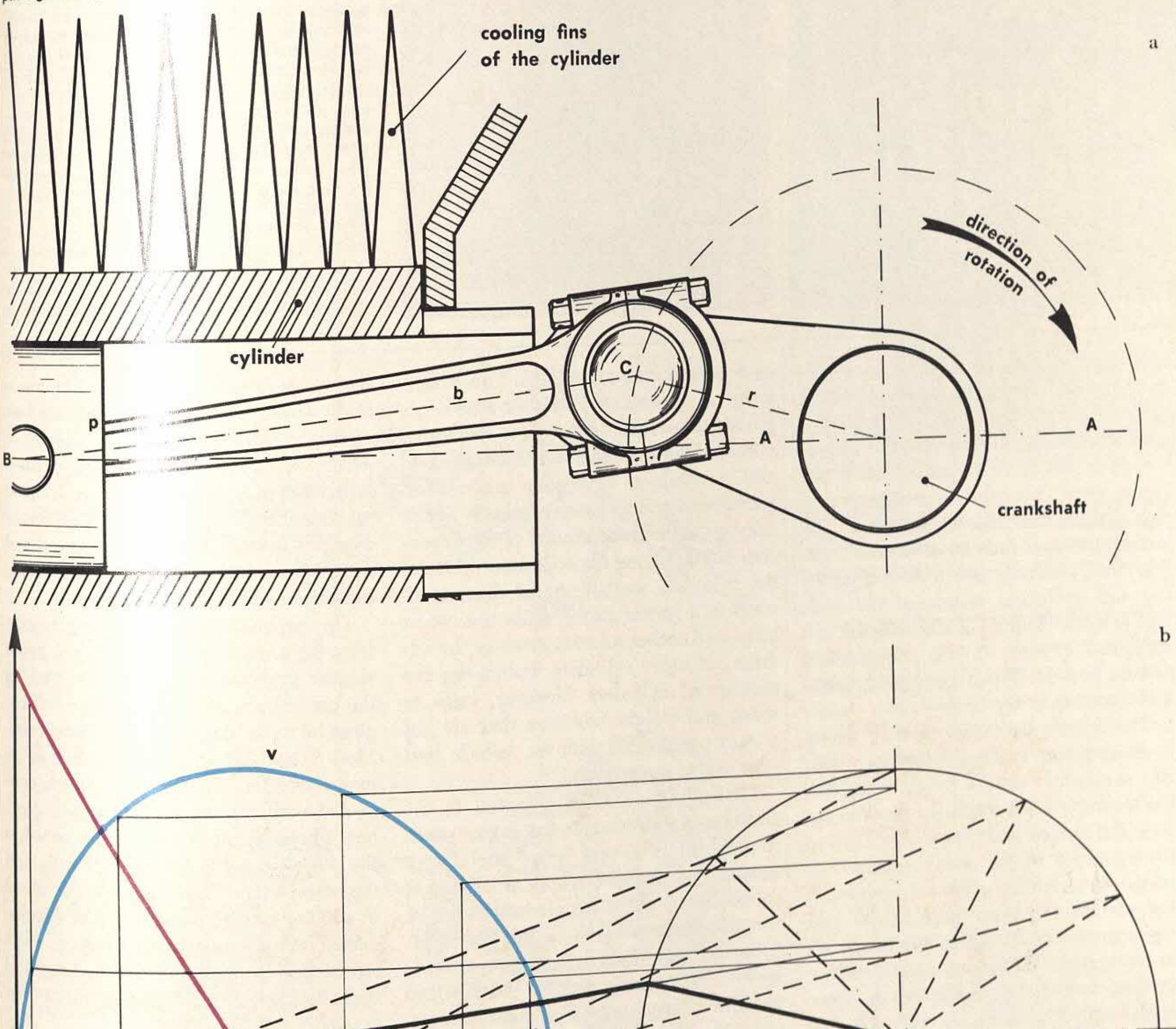
A CRANKSHAFT—This photograph shows the crankshaft of an automobile engine. The crankpin bearings **b** and the main bearings **m** are clearly shown. The counterweights **c** partially counteract the forces of inertia that come into play during rotation of the shaft. These weights reduce the task performed by the flywheel and contribute to the smooth running of the engine.



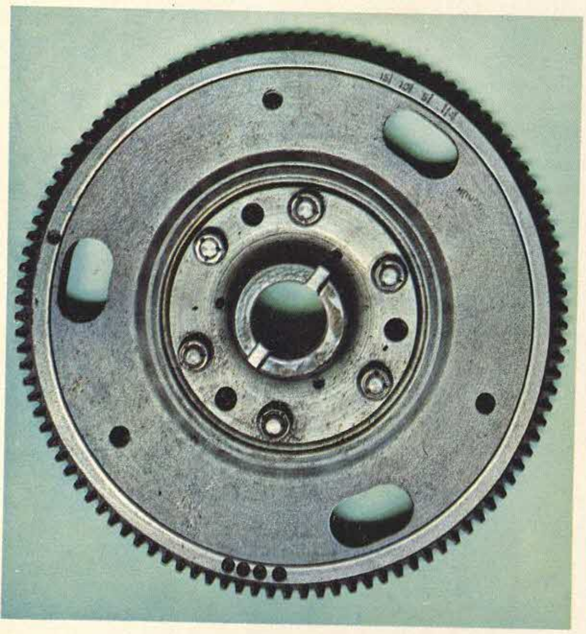
3

THE CRANKSHAFT-CONNECTING ROD SYSTEM—Illustration 3a is a schematic diagram of the cylinder of an internal combustion engine, together with a piston **p** and a connecting rod **b,** to which the piston is connected. The connecting rod in turn is attached to the engine shaft by means of a pivot that is eccentric with respect to the axis. That is, it is connected to the shaft by means of a crank. All points on the crank follow a circular path during the rotation of the crankshaft, while the various points of the connecting rod pass gradually from rectilinear motion to circular motion. The small end of the connecting rod **B,** which is attached to the piston, executes rectilinear motion along the axis **AA,** while the other end of the connecting rod, the large end, is connected to the crank pin **C.** The path described by the large end of the rod is represented by a circle having the radius **r.**

In Illustration 3b, the connecting rod and the crank are represented by line diagrams. The segment **BB′** represents the route traveled by the piston. If the motion transmitted to the crankshaft is to be uniform (the crankshaft must rotate at a uniform speed), the piston in the cylinder must travel along the segment **BB′** with well-defined speeds and accelerations. These are easily determined by means of a graphic construction—which, for the sake of brevity, will not be described here. The two curves **v** and **a** represent, respectively, the speed and the acceleration of the small end of the connecting rod; that is, the speed and acceleration of the piston in the cylinder. The curve **a** also represents the forces of inertia that come into play during the motion (when suitably scaled).



4a



4b

A FLYWHEEL AND A MODEL OF THE CRANKSHAFT-CONNECTING ROD SYSTEM—The flywheel shown in Illustration 4a is a part of an automobile engine. Similar devices may be used in other mechanisms to regulate rotational motion and to produce smooth running conditions. Given its large size and mass, the flywheel has a substantial moment of inertia. Hence, once it has been put in motion, it tends to maintain a smooth and regular motion of the system. In a four-stroke internal combustion engine, only one stroke out of every four made by a piston is a power stroke; therefore, if the engine were not equipped with a flywheel, it would have an unacceptably jerky motion. Illustration 4b shows a demonstration model of the crankshaft-connecting rod system.

The easiest way to connect the piston to the engine shaft is to use a connecting rod and a crankshaft (see Illustration 3). The connecting rod couples two mechanical parts that move in fundamentally different ways. One exhibits rectilinear motion and the other rotational motion. The various points of the connecting rod must, therefore, gradually pass from reciprocating and rectilinear motion at the small end to a combination of translational and rotational motion at the intermediate points, and finally to completely rotational motion at the large end.

The system consisting of the piston, connecting rod, and crankshaft has mass; this means that inertial forces occur during its motion. These forces tend to oppose the motion and create vibrations in the operation of the entire engine. The piston, because it performs a reciprocating motion and must periodically pass from motion in one direction to motion in the opposite direction, is most subject to these inertial forces. During the stops at the end of each stroke, or rather immediately before and immediately after these stops, deceleration and acceleration respectively reach their maximum values. It is at these moments that the inertia force acting on the piston attains its peak.

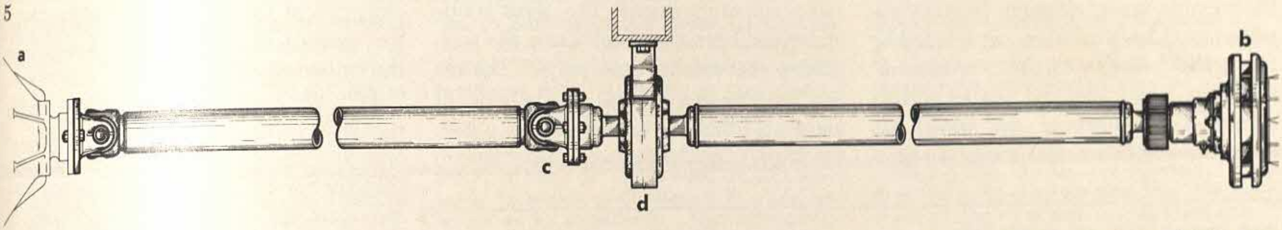
Therefore, the motion of the piston becomes discontinuous or jerky. This irregularity is accentuated in four-stroke engines that perform a complete revolution without load (exhaust and intake strokes) and a complete revolution under load (compression and power strokes). A first solution to the discontinuity problem consists of staggering the operation cycles of the cylinders so that one of them is always in a power stroke while the others are not. Another solution involves the addition of more cylinders. Increasing the number of cylinders, however, leads to space and weight problems that are not always compatible with the vehicle that is to utilize the engine.

The solution usually adopted is the addition of a flywheel to the engine shaft. A flywheel is a very heavy metal disk constructed so that the bulk of its mass is concentrated along the circumference of the wheel. As a result of this mass distribution, a flywheel has a large rotational inertia. This inertia permits the flywheel to store a large amount of rotational energy. By releasing this energy during the time when the inertia forces on the pistons, connecting rods, and camshaft are large, the effect of these forces on the engine output is minimized. During the time when the inertia forces are small, the flywheel stores energy. Another solution—used in conjunction with a flywheel—is the use of counterweights on the crankshaft. These weights are either fused on or forged integrally with the crankshaft itself. Such counterweights are particularly large in engines that have a limited number of cylinders; they compensate the flywheel, when, because of space and weight considerations, its size cannot be increased beyond a certain limit.

The problem of transmitting motion from the engine to the wheels is a much simpler problem; it can be resolved by the use of belts or chains, as in small engines for industrial use, or on motorcycles. More often, however, and especially in motor vehicles, a steel drive shaft is connected with the engine through a gearbox. The drive shaft is then connected to the wheels via the differential gear and the wheel axles. This system is described in another article; therefore, it is not discussed here. Certain parts, however, form an integral part of the drive shaft, and these flexible joints do require further discussion.

A motor vehicle, regardless of its size, is not a rigid system. Its parts are not rigidly welded together so that relative movement between parts is impossible. Just the opposite is the case; the chassis is very elastic. Furthermore, the engine is not rigidly fixed to the chassis; rather, it is mounted on flexible rubber supports

5

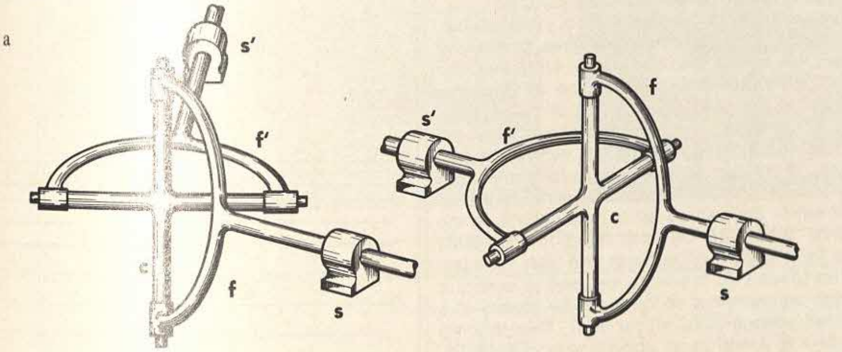


THE DRIVE SHAFT—This illustration is a schematic drawing of an automobile drive shaft. This unit joins the engine and gearbox unit **a** to the differential **b**. The drive shaft consists of a steel tube divided into two roughly equal parts by the universal joint **c** and the flexible joint **d.** These joints permit small alignment changes of the drive shaft. They are needed because the wheels cannot be rigidly connected to the engine. The presence of these joints also ensures smooth motion.

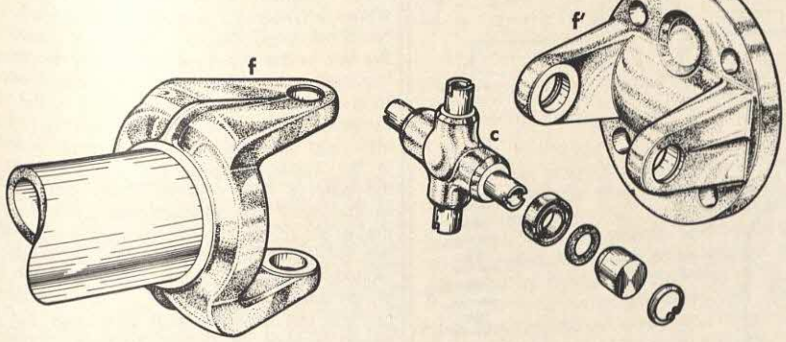
6

THE UNIVERSAL JOINT—Illustration 6a is a schematic representation of a universal joint. Motion is transmitted from one part of the shaft to the other via an intermediate element **c**, which has the form of a cross. The two parts of the drive shaft are connected to this cross by means of the yokes **f** and **f'**, and each yoke is hinged to the two ends of one arm of the cross as shown in the diagram. If the two halves of the shaft are fixed at one point each by means of the bearings **s** and **s'** and then rotated, the rotation is transmitted smoothly from one to the other, even when the two parts are misaligned. Illustration 6b, on the other hand, shows a universal joint of the type usually used with automobile drive shafts. This type of joint is far more compact than the one shown in Illustration 6a, although its two essential elements, the cross and the yokes, are easily recognizable.

a



b



7

THE CROSS OF A FLEXIBLE UNIVERSAL JOINT—This is the cross of a special type of universal joint. In this case the cross is made of rubber. The joint produced is also flexible.

that dampen vibrations. As a result, these vibrations are not transmitted to the rest of the vehicle.

Similar conditions apply to the wheels, which are connected to the chassis by means of a flexible suspension system consisting of shock absorbers, springs, floating arms, and other parts. The rear wheels and axles, therefore, have considerable freedom of movement relative to the car as a whole and to the engine in particular.

If the drive shaft were rigidly connected to the aforementioned parts, it would be under stress, would operate poorly, and would run the risk of fracturing, thus threatening the integrity of other parts. Flexible joints, therefore, provide the necessary degree of freedom between moving parts. This freedom of movement, however, does not interfere with power transmission.

Although this article has been restricted to one type of engine and vehicle, the topic is of broader importance. In every technical problem involving the transformation and transmission of energy, similar difficulties are encountered.

TRIODES | simple voltage amplifiers

The triode is the most important of all the vacuum tubes, although in many applications it has now been superseded by transistors because of their smaller dimensions, longer life, and smaller current consumption. However, the triode has some characteristics that make it necessary in certain applications, especially in radio communications. The term triode is derived from the Greek (as is the term diode) and means "three ways." The triode has three electrodes: the cathode, or filament, identical to that of the diode; the plate, also identical in its functions to that of the diode; and the grid, the third electrode of the triode. The grid controls the amount of current that passes from the cathode to the plate.

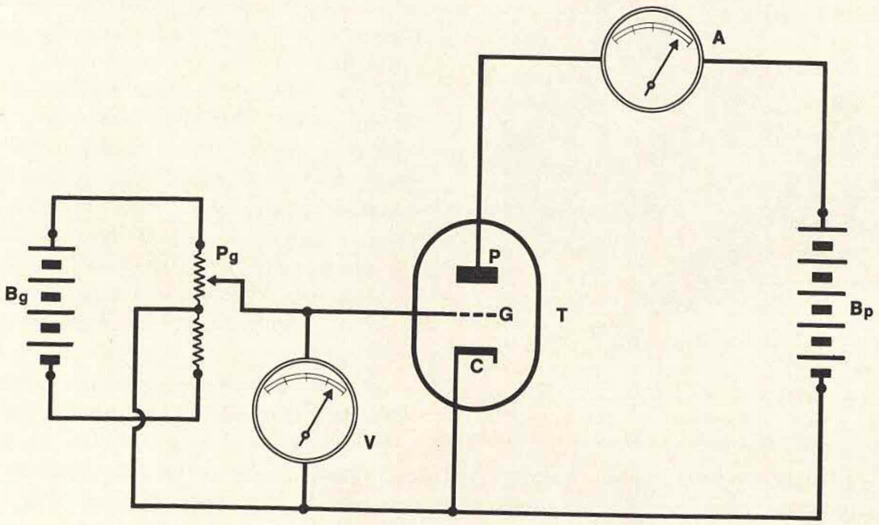
The operating principle of the triode is relatively simple. Because its applications are based on its ability to amplify the grid current, a brief comparison with the operation of the diode may serve to show the triode's special function. When a diode is switched on, a current flows between the cathode and the plate. If the plate voltage is constant, the plate current will also be constant. In a triode, on the other hand, if the plate voltage

1

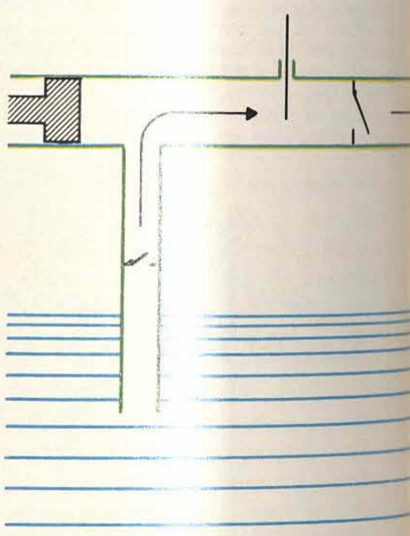
THE TRIODE AND ITS INPUT—This illustration shows the operation of a triode. The triode as a whole is indicated by the letter **T,** and consists of the cathode **C,** the grid **G,** and the plate, or anode, **P.** Only that part of the cathode that actually emits the electrons is represented here. The filament and the input circuit that make it incandescent are not shown. The plate circuit consists of a voltage generator (symbolically represented here by the plate battery $\mathbf{B_p}$). Its negative pole is connected to the cathode and its positive pole to the plate. Also included in the plate circuit is an ammeter **A,** which measures the current flowing in the circuit.

Although, for the sake of simplicity, the plate is shown as being supplied with voltages that are positive with respect to the cathode (the tube will thus always be in conditions analogous to those of a diode that is conducting), the grid has been connected to a circuit that permits it to become positive or negative with respect to the cathode. This circuit consists of a battery whose poles are connected to the terminals of a potentiometer. The battery is indicated by the letters $\mathbf{B_g}$ and the potentiometer by the letters $\mathbf{P_g}$. The arrow pointing toward the potentiometer serves to obtain a voltage that, according to the position of the arrow, can be either positive or negative with respect to the cathode. A voltmeter **V** serves to measure the grid voltage.

The operation of the triode may now be explained in a simple manner. Suppose that the grid (which in reality is only a thin wire mesh around the cathode) does not exist. When a voltage is applied to the plate circuit, the triode will behave like a diode and a current will flow through it—that is, the electrons emitted by the cathode (being negative) will be collected by the positive anode placed opposite. In this condition, the motion of the electrons occurs exactly as it does in the diode. Now suppose that a variable voltage is applied to the grid. Initially, this voltage is negative with respect to the cathode. If it is extremely negative, it will ensure that the electrons emitted by the cathode will be repelled backward, thus neutralizing the positive field of the plate. However, if the grid voltage becomes less negative with respect to the cathode, the electrons will no longer be subject to a strong repelling action and, at a certain value of the grid voltage, they will succeed in passing beyond the grid and being collected by the plate. In this condition, current will begin to flow in the plate circuit. If the grid voltage now becomes positive, it will not hinder the flow of the electrons and will even promote the migration of the electrons to the plate (as a result of the positive grid potential, a part of these electrons are collected by the grid itself). A very small current, therefore, flows in the grid circuit. Because the grid voltage is also small, the total power in the grid circuit is rather small. In the plate circuit, on the other hand, it is possible to have fairly high currents and high voltages. The triode can, therefore, be used to control a high voltage (that of the plate) by means of a small voltage (that of the grid). This is a triode's function as an amplifying and modulating device



2



A HYDRAULIC ANALOGY — This illustration shows a hydraulic analogy of the triode. The horizontal pipe draws in water. A piston (on the left in the pipe) moves with a reciprocating motion; it therefore causes water obtained from the basin below (represented by the blue horizontal lines) to flow up through the vertical pipe and into the horizontal pipe. A valve in the horizontal pipe permits the water in the pipe to flow only from left to right. Up to this point the flow scheme is identical to that of the diode; the one-way valve represents the rectifying action of the triode, which is analogous to that of the diode. Between the piston and the one-way valve is a gate valve (on the top of the pipe); this valve can be raised or lowered, thus permitting or preventing the flow of water through the pipe. The flow of a large quantity of water can be controlled or modified by means of the small power that is required to operate the gate valve. As can be seen, the analogy between the triode and this hydraulic system is very close indeed.

3

TRIODES FOR A BROAD RANGE OF POWERS —Illustration 3a shows a nuvistor. This is a small triode used in electronic circuits in which the power to be controlled amounts to a few watts. It is made of ceramic and is capable of withstanding temperatures in excess of 100° C (212° F). Some models can withstand temperatures of several hundred degrees as well as strong vibrations. This type of device also has a very low power consumption. In ordinary vacuum tubes, on the other hand, the power consumption is always high because the heating of the filament requires a large quantity of electric energy, most of which is dissipated in the form of heat.

Illustration 3b shows a triode suitable for powers that are not excessively high, for example, of the order of 100 watts. A high-power triode is shown in Illustration 3c. This type serves as an amplifier for radio broadcasting and has applications in control systems.

Powers of the order of many kilowatts can be easily controlled with triodes. High-powered triodes must be cooled in order to dissipate the heat emissions from the filament, as excessive heat could be harmful and render the triode inefficient. The filaments of low- and medium-power triodes are generally made from thorium-plated tungsten, but the high-power triodes have pure tungsten filaments.

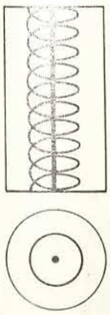
a

b

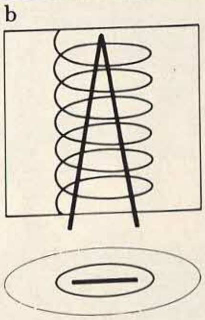
c

4

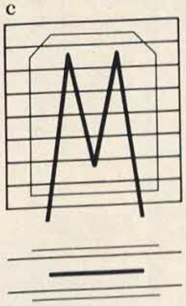
a



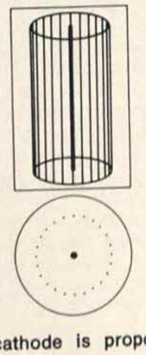
b



c



d



ELECTRODE CONFIGURATIONS — These four illustrations show the relative positions of the cathode, grid, and plate in some typical triodes. The grid consists of a mesh of extremely thin wires that are quite widely spaced, thus allowing the free movement of the electrons and also preventing the electrons from being deposited on the grid when it is positive with respect to the cathode. This is a real advantage because a small grid current will ensure a high power amplification factor. As shown in all the illustrations, the grid is always placed between the cathode and the anode. This arrangement of the electrodes shows how the current is conducted through the triode. The simplest case is that where the electrodes of the tube have a symmetric shape—the cases shown in Illustrations 4a and 4d, for example. In these examples, both the electric field in the vicinity of the cathode and the plate current are determined by three characteristics of the tube. Two of these characteristics are variable. They are, respectively, the plate voltage $\mathbf{V_p}$ and the grid voltage $\mathbf{V_g}$. The third characteristic is one that depends only on the shape of the tube; it is called the amplification factor and is denoted by the Greek letter μ. More precisely, it can be demonstrated that the voltage in the proximity of the cathode is given by $\mathbf{V_g + V_p/\mu}$; moreover, the current that issues from the cathode is proportional to this voltage raised to the power 1.5: $\mathbf{I_p = K(V_g + V_p/\mu)^{3/2}}$. In case the grid voltage is negative, all of this current flows toward the plate; no electrons become deposited when the grid is negative with respect to the cathode. On the other hand, when the grid is positive with respect to the cathode, some of the electrons will become deposited on the cathode and some on the grid. This means that the plate will receive a smaller quantity of electrons than is required by the formula. Nevertheless, in view of the shape of the electrodes, the plate will always receive most of them.

is kept constant while the voltage of the grid is varied, the grid voltage will affect the flow of the electrons from the cathode to the plate and cause a variation of the current that reaches the plate.

In the case of the triode, the grid might be described as having a function similar to that of a faucet—the opening and closing of this "faucet" causes variations in the current flowing through it. The triode is sensitive to changes of less than a millionth of a volt.

5

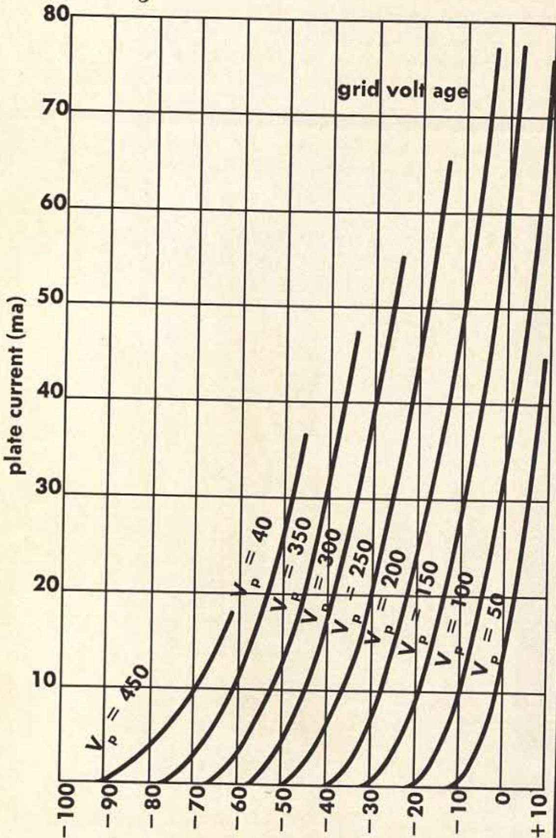
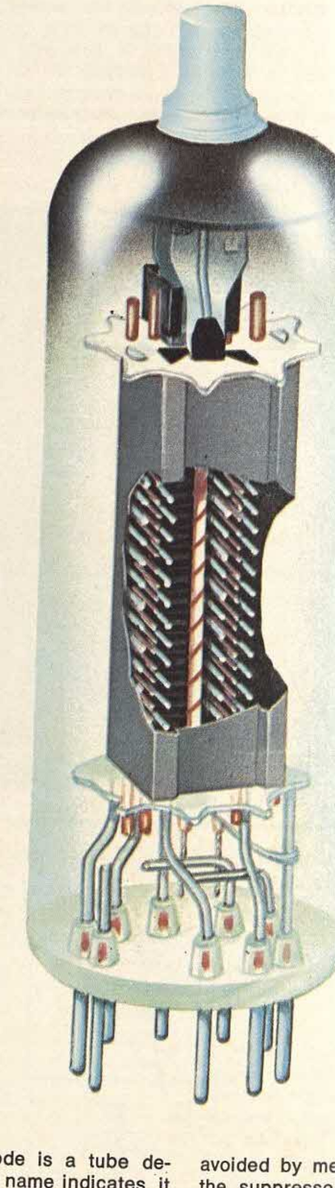


PLATE CHARACTERISTICS — The characteristics of a triode have to be known before it can be properly used in a circuit. The term plate characteristics is understood as referring to the curve that gives the intensity of the plate current in terms of the grid voltage. Such a curve can be drawn for various values of the plate voltage, and the complete diagram of the plate characteristics therefore contains many curves. This illustration shows such a family of curves. The curves toward the left are those concerning situations where the plate voltage is at or near its peak value. In fact, in this case, it would be necessary to select a strongly negative grid voltage in order to neutralize the effect of the very strong plate field. For low plate voltages, on the other hand, it is quite sufficient to have a rather low grid voltage in order to ensure that the plate will receive a strong current. A very characteristic and important feature of a triode is represented by the interdiction potential, which is the voltage that must be applied to the grid in order to prevent the tube from conducting. For each plate voltage there will be a particular value of the grid voltage that will reduce the plate current to zero. Quite obviously, the greater the plate voltage the greater will have to be the negative voltage that must be applied to the grid in order to obtain complete interdiction; that is, to prevent the electrons emitted by the cathode from reaching the plate. To prevent the tube from conducting, $V_g + V_p / \mu$ must equal zero; this means that V_g must be equal to $-V_p / \mu$.

6



THE PENTODE—The pentode is a tube derived from the triode. As its name indicates, it consists of five electrodes. These are the cathode, the control grid, and the plate—all of which are analogous to the corresponding electrodes of an ordinary triode—as well as two additional grids placed between the control grid and the plate. The first of these grids, the one nearer the control grid, is known as the screen; the other, closer to the plate, is called the suppressor grid. The functions of the first three electrodes are identical to the functions they perform in the triode. The fourth, the screen, screens the cathode from the plate potential; by applying an appropriate voltage to the screen grid, conditions are created in which the plate potential will be negligible at the surface of the cathode. In spite of this, it might still be possible for some electrons to reach the plate; this can be avoided by means of the suppressor. In fact, the suppressor potential can stop the electrons in the proximity of its internal surface (the one facing the cathode); here the electron cloud forms a space that is densely filled with electrons, which may be extracted by means of an appropriate plate potential. This electron-filled space, therefore, fulfills the function of a cathode and is often called the virtual cathode.

The movement of the electrons in the pentode can be much better controlled than in the triode. Nevertheless, its function is identical to that of the triode; in fact, the pentode acts as an amplifier and a modulator. A somewhat simpler tube than the pentode is represented by the tetrode. This valve has only four electrodes and could be described as a pentode without a suppressor grid. A tetrode, therefore, has exactly the same functions as a triode.

ULTRAVIOLET RAYS | properties and effects

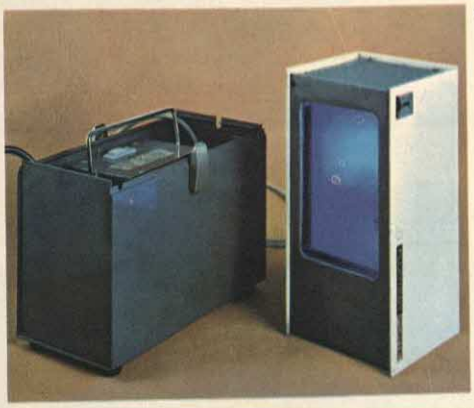
Invisible radiations have many interesting uses. The study of such radiations not only demonstrates how much the human eye does not see; it also increases what man knows about the structure of the atom. Spectroscopic analysis of visible radiation permits the interpretation of phenomena concerning the outer electrons of atomic structure. Radiations beyond the visible spectrum can reveal the internal structure of the atom itself. For example, ionic changes are accentuated by ultraviolet radiations, and infrared radiation is used for the study of molecular structures.

One difficulty in using invisible ràdiations is the need for special equipment. While no particularly special equipment is needed for working with visible light rays, special sources and detection instruments are required for the employment of infrared, ultraviolet, x-rays, and radio waves. The cost of such equipment varies, however, with the different types of radiations. Infrared requires much more costly equipment than does ultraviolet. The experimenter, for example, can use photographic plates with ultraviolet because its rays show up clearly on them, while infrared rays show up on the plates only in the shorter wavelengths near the limit of visible light. Ultraviolet radiations bring about the phenomenon of fluorescence, while infrared rays produce no similar effect.

Photographers have found many uses for ultraviolet rays. The rays do not penetrate glass, but a quartz lens permits them to pass. Pictures can be taken in the dark by turning radiation from an ultraviolet lamp on the object.

Ultraviolet fluorescence is occasionally used to detect forged documents; the different kinds of ink on the documents can be distinguished by their fluorescence. Textile manufacturers use ultraviolet rays to distinguish one material from another. Gems and metals are often recognized

2



a

b

ULTRAVIOLET SOURCES—This illustration shows two common sources of ultraviolet radiation: one is used when intense radiation is needed with moderately short wavelengths; the other is used to produce a line spectrum. The first (Illustration 2a) is called Wood's light, after the American physicist Robert Wood, who carried out many experiments with ultraviolet radiation. Ultraviolet and visible radiations are produced in this light by the agitation of mercury vapors inside the bulb. In making experiments, however, it is preferable to produce pure ultraviolet radiation and not a mixture of ultraviolet and visible rays. For this reason, it was necessary to devise a way in which visible rays were prevented from leaving the light source. The glass screen devised by Wood does just that—it stops visible radiation but lets ultraviolet rays pass through. The light, observed in a dark room, gives off a faint violet color.

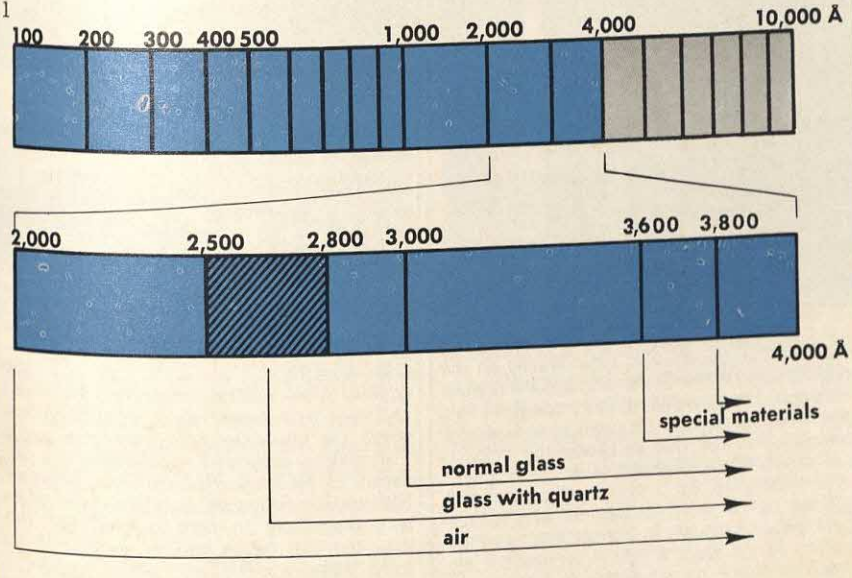
The second source of ultraviolet rays (Illustration 2b) is a mercury vapor light in a quartz bulb. It is important that the mercury be at a low—but not too low—pressure. The pressure ensures an intense radiation as well as the line spectrum. Intensity of radiation is reduced if the mercury pressure is lowered. This is significant in certain spectroscopy experiments.

THE ULTRAVIOLET FIELD—The ultraviolet spectrum is divided into various zones. The top band shows the entire ultraviolet spectrum. It extends from 100 Å to 4,000 Å. Below 100 Å the spectrum extends into the area of x-ray radiation. Above 4,000 Å (shown in gray) is the region of visible radiation; beyond that, the spectrum goes into the infrared zone. It is costly and difficult to work in the area between 100 Å and 2,000 Å because vacuum conditions are required to employ ultraviolet radiation for experiments in this range. The examples shown are, therefore, in the 2,000 Å to 4,000 Å range.

The lower band shows the finer distinctions within the 2,000-4,000 Å range. Between 3,800 Å and 4,000 Å the human eye can see and distinguish the colors of ultraviolet rays. For wavelengths of 3,600 Å and above, normal glass can be used for small, thin lenses, while special materials must be used for shorter wavelengths. Down to 3,000 Å, high-quartz-content glass is suitable for lenses; below 3,000 Å, either pure quartz or high-quartz-content glass of less than 1 mm in thickness must be used.

At about the 2,800 Å to 2,500 Å limit, the eye no longer sees any radiation. Down to this limit, however, intense ultraviolet radiation creates a sensation of grayish light. Below the 2,800 Å-2,500 Å limit, the interior of the eyeball becomes fluorescent under the stimulus of ultraviolet radiation. Below 2,000 Å, the transmission of ultraviolet rays in air is no longer possible.

1



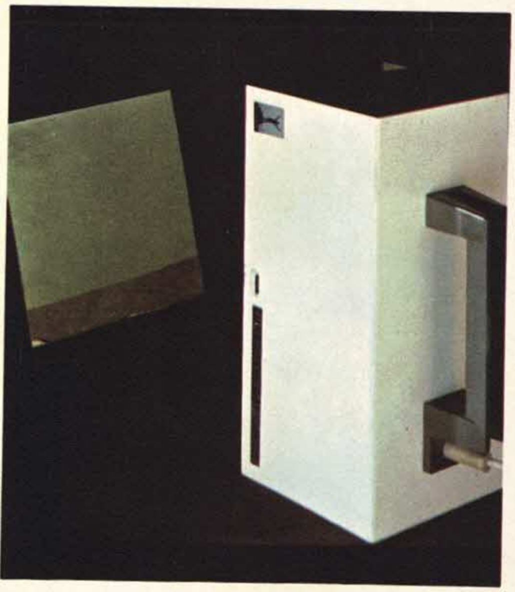
by their ultraviolet fluorescence. Dentists can distinguish a live tooth from a dead one because live teeth fluoresce, whereas dead ones do not.

Ultraviolet radiation from sun lamps can give certain benefits; the rays act like sunlight to make the skin secrete the dark pigment that tans the skin. The rays also change ergosterol under the skin into Vitamin D.

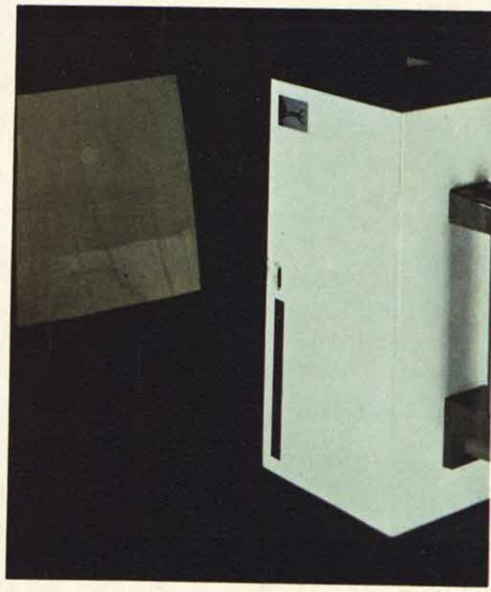
When used to illuminate specimens observed under a microscope, ultraviolet rays provide two chief advantages over visible light. First, resolution can be increased because of the shorter wavelength, with the result that finer details

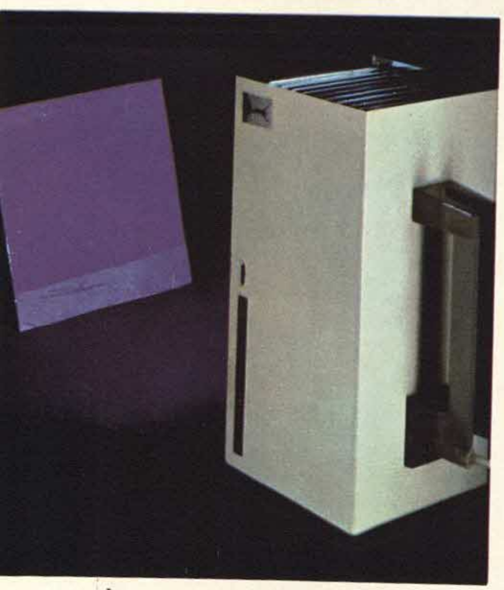
3

a

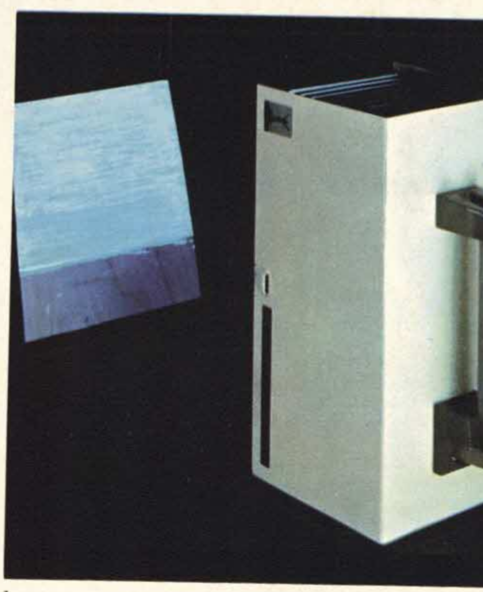


c





b



d

FLUORESCENT MATERIALS — In order to show the presence of ultraviolet radiation it is necessary to use fluorescent materials such as those used for advertising billboards. In Illustration 3a a piece of cardboard painted with some of this material is illuminated with normal light: it appears its own green color. In Illustration 3b, illuminated by a Wood's light, it appears violet and a violet luminescence is present around it.

The violet luminescence derives from the fact that the intense radiation sent into the room by the Wood's light acts directly on the photographic plate. Developing solutions show ultraviolet as a violet-colored radiation. This does not mean that rays are violet-colored, because the eye that perceives the color is not sensitive to ultraviolet rays.

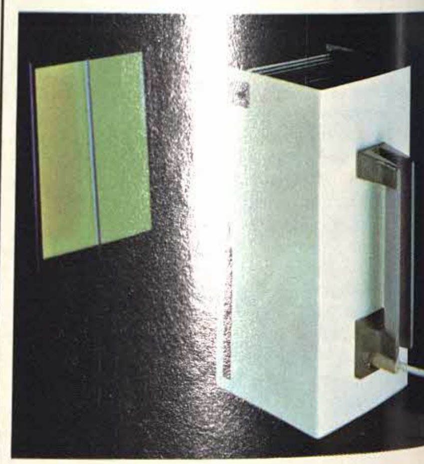
In Illustration 3c a thin sheet of plywood, the top part of which is covered with fluorescent paint, is shown in normal light. In Illustration 3d the Wood's light is on and the plywood is struck by ultraviolet rays that bring out the fluorescence—in this case, blue.

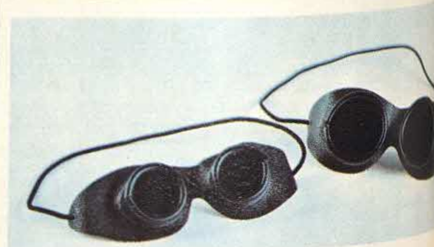
4

OPAQUE AND TRANSPARENT MATERIALS— It is easy to study the transparency of certain materials to ultraviolet rays using Wood's light. An excellent instrument can be built for this purpose by simply putting Wood's light behind a screen (or in a box) with a slot opening so that only a narrow beam of ultraviolet light can escape. A screen to intercept ultraviolet rays is then placed behind the object whose transparency is to be measured. Measuring the transparency of an object is accomplished by measuring the degree to which it absorbs radiation (since absorption and transparency are reciprocal values). In the first photograph (Illustration 4a) there is a thick sheet of glass behind which the fluorescence is much weaker; this means that the

a

b





c

glass absorbs the ultraviolet light to a large degree. In the second photograph (Illustration 4b) there is a thinner sheet of glass through which the ultraviolet rays pass more easily.

In making ultraviolet experiments it is necessary to take some precautions. Ultraviolet light causes erythema, or inflammation of the skin (especially in light-complexioned people) and can cause serious damage to the eyes. While working with ultraviolet light, it is wise to wear special glasses such as those shown in Illustration 4c.

5

CHEMICAL ACTIVITY—Chemically, ultraviolet light is extremely active. One way of showing the effect of ultraviolet rays on certain substances is to demonstrate how it can bleach a piece of cloth. Ultraviolet rays are particularly effective in destroying complex-structured molecules, such as those in dyes. To show this effect, a colored sample (cloth or paper) is exposed to ultraviolet light. In Illustration 5a the central part of the cloth has been covered with a metal strip. After exposure to ultraviolet radiation, the exposed part of the sample appears lightened or bleached when it is compared with the part that was covered (Illustration 5b).

a

b

6

GERMICIDAL POWER—Another characteristic of ultraviolet light is its germicidal power. This property is related to the preceding one: living structures undergo the effect of the same molecular destruction by ultraviolet radiation as the destruction that causes bleaching. The effect is more pronounced in smaller organisms whose entire bodies can be subjected to radiation at the same time. Two bacteria cultures in Petri dishes are shown in Illustration 6a. One culture is protected from ultraviolet light; the other is placed under intense radiation. Illustration 6b shows the result of the experiment: while the bacteria colonies in the right dish have flourished and developed to several millimeters in size, the left dish is free of bacteria. Its bacteria were unable to form colonies because the strong germicidal power of ultraviolet rays destroyed the bacteria culture.

The experiment described here is partial and limited; in fact, the rays from a Wood's light have wavelengths longer than 3,000 Å. Therefore, these are results of an ultraviolet light band very near visible light. If a shorter wavelength had been used, the culture bacteria would have been killed very rapidly. For this reason, germicidal ultraviolet lights are constructed to give ultraviolet rays of about 2,500 Å.

a

b

can be seen than with visible light. Second, biological materials display increased differential absorption to portions of the ultraviolet spectrum, a characteristic of importance to the biochemist and the bacteriologist. Ultraviolet radiation and quartz monochromatic objectives corrected for only one ultraviolet wavelength are employed.

Three important operations are required to work successfully with ultraviolet radiations. These consist of: producing the radiation; making an apparatus to reflect, converge, or diverge the ultraviolet rays; and rendering the ultraviolet rays visible by means of fluorescence or the use of photographic solutions.

In working with ultraviolet radiations, their wavelength is an important factor. The nearer the ultraviolet radiation is to visible light, the less is the difficulty in exploiting its properties. In working with radiations having wavelengths less than 2,000 Ångstroms, air can no longer be considered a transparent media; such an experiment must be made in a vacuum, or at least in the absence of oxygen, which is the main component of air responsible for the absorption.

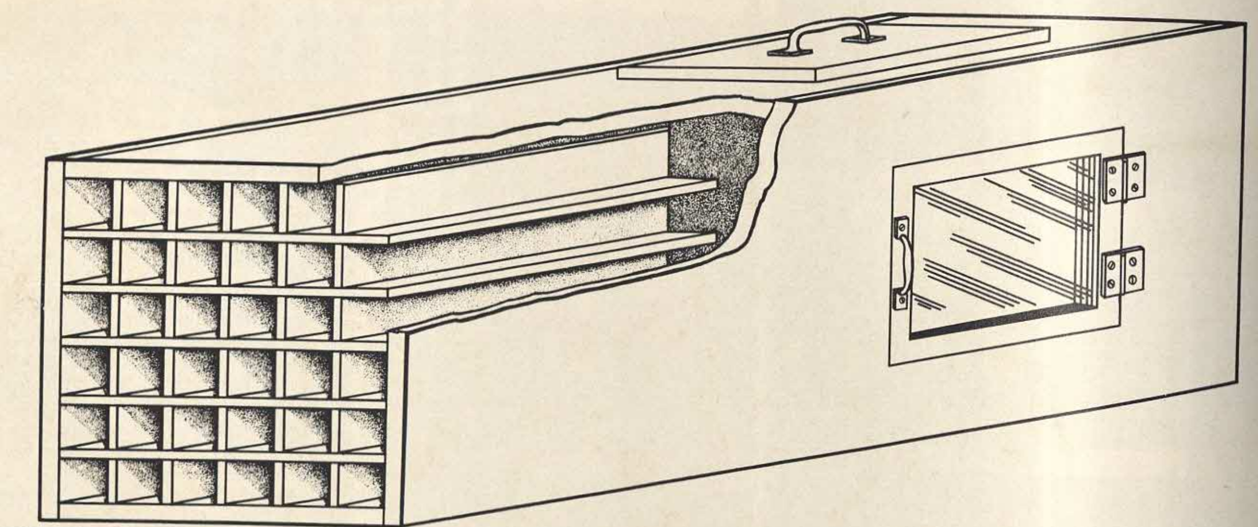
THE WIND TUNNEL | a useful tool for studying aerodynamics

Aircraft design owes a great deal to the wind tunnel, a structure within which a stream of moving air can be carefully controlled. When a stationary object is suspended in rapidly moving air, the effect is almost exactly the same as when the object moves rapidly through stationary air, as when an aircraft flies. Entire aircraft, or parts of them, are mounted inside the wind tunnel, where sensitive instruments measure accurately the forces that act on different parts of the aircraft's surfaces as air moves past it. Engineers who study the results of wind tunnel tests can predict accurately how an aircraft will perform in actual flight.

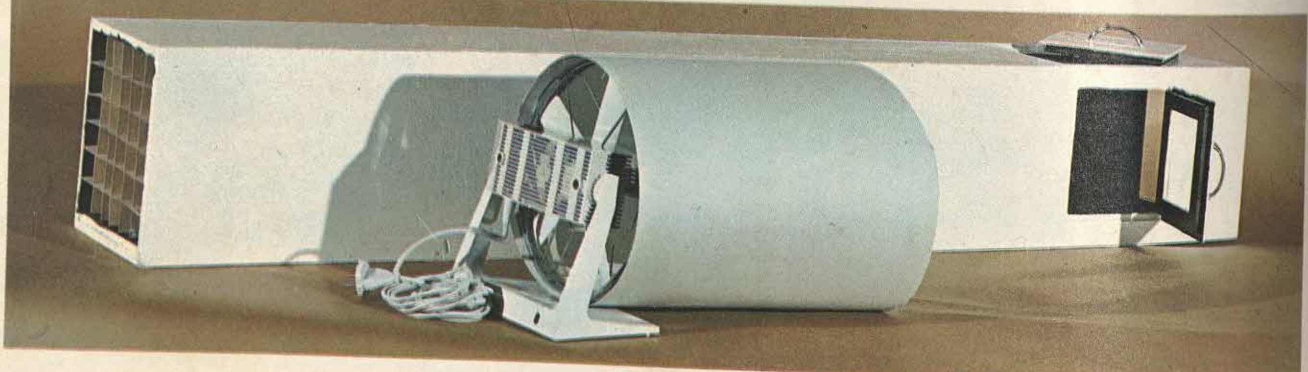
Wind tunnels large enough to accommodate full-size aircraft are expensive to build and equip. Fortunately, for most purposes, a model aircraft mounted inside a miniature wind tunnel serves just as well.

Spacecraft, rockets, and missiles of all kinds are tested in wind tunnels, which

1a



1b



BUILDING A SMALL WIND TUNNEL—It is fairly easy to build a model wind tunnel that can be used to investigate a wide range of aerodynamic phenomena. If the construction is to be kept simple, it must be an open circuit design, with air moving at subsonic speed. An ordinary electric fan, such as is used for ventilation, can provide the stream of air.

It is not much more difficult to build a small-scale supersonic wind tunnel, but there would be little point in doing so. It would be difficult to observe and measure the effects of high-speed winds without special instruments that are expensive and not readily available.

The wind tunnel shown in this illustration is essentially an elongated box with a square cross section 20 cm x 20 cm (about 8 in. x 8 in.). The length is 1.5 m (about 5 ft). The cross section might well be rectangular, with the longer dimension horizontal, if model aircraft with long wings are to be accommodated. The box is built of wood and painted black inside. The basic parts should be assembled in such a way that the inside surface is smooth; any extra pieces of wood needed to reinforce the structure should be attached on the outside rather than on the inside.

Only two parts of the structure are at all complicated. One is the duct that channels air from the fan into the tunnel. If the fan is round and the wind tunnel square, the duct will need to be a flexible sleeve made of plastic or cardboard. The other relatively complex structure is the "eggcrate" or honeycomb grill that fills in the 30 cm (about 12 in.) of the tunnel nearest the fan. This part can be constructed of very thin plywood or fiberboard, sheet aluminum, or some other easy-to-cut but rigid material. It serves to establish a uniform flow of air, eliminating any twisting motion (vorticity) imparted by the whirling blades of the fan.

Near the end of the tunnel that is farthest from the fan is an observation window, which can be mounted in a hinged frame, and another opening, which should have a lift-off lid. Both lid and window should be constructed so that when in place, the inside surface of the tunnel is flat and smooth, without cracks or ridges that would disturb the flow of air.

Illustration 1a shows construction details; Illustration 1b is a photograph of a completed wind tunnel with window and access door open with the fan unattached.

are also used to evaluate the streamlining of automobile bodies and trains, and to test the effects of high winds on such structures as tall buildings and bridges.

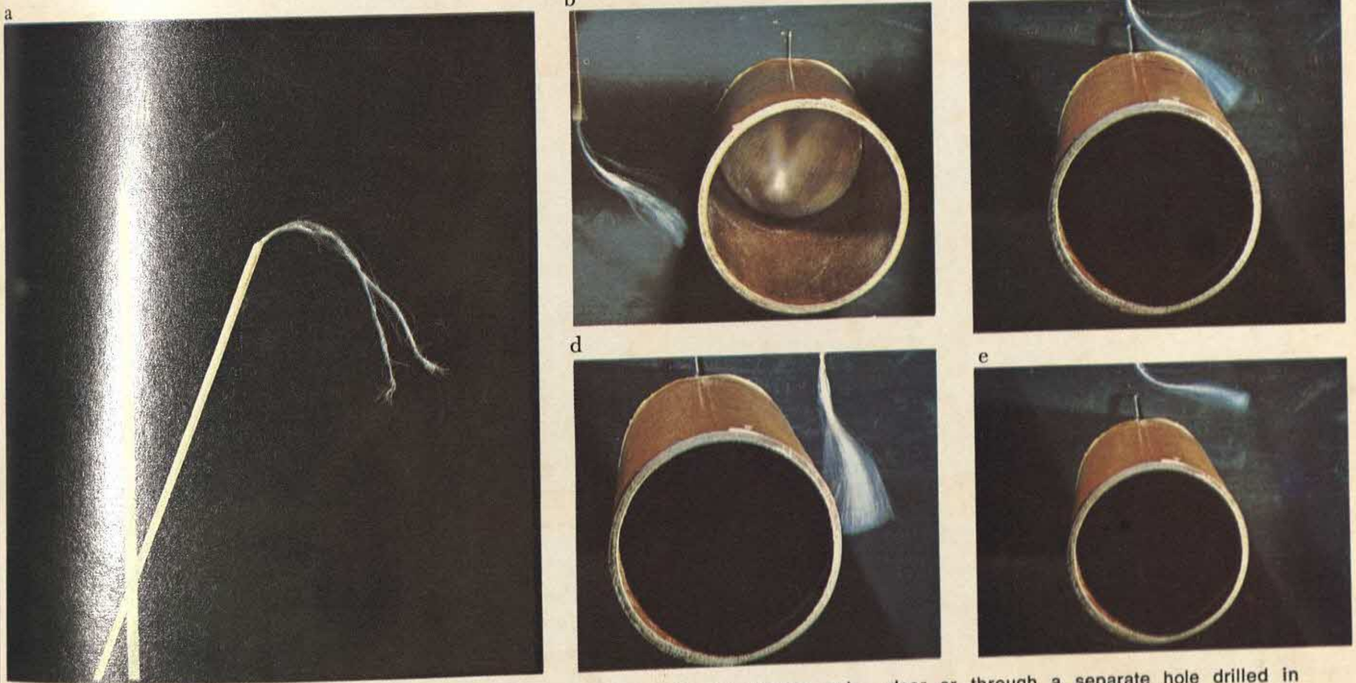
Wind tunnels are classified according to the speed with which the air moves inside them. In low-speed or subsonic wind tunnels the air moves at less than the speed of sound. In these installations, large electric fans move the air through the tunnel. In high-speed or supersonic wind tunnels, giant air compressors provide a blast of air that moves faster than the speed of sound.

Wind tunnels are also classified as open circuit or closed circuit designs. In the former, both ends of the tunnel are open; air is drawn in from the surrounding space at one end and released at the other. In the latter, a return duct connects the two ends, so that air inside the tunnel recirculates. In a closed circuit installation, the temperature and air pressure inside the tunnel can be controlled to simulate special conditions such as flight at high altitudes.

Little control can be maintained over the pressure, temperature, and humidity of the air in the open circuit model because the ends of the tunnel circuit are exposed to the atmosphere. Therefore, the most common of all wind tunnel designs is the closed circuit tunnel.

Through its simulation of the conditions of very high-speed flight, the wind tunnel has enabled engineers to examine problems associated with travel both inside and outside the atmosphere of the Earth.

2



MAKING THE FLOW OF AIR VISIBLE—Professional engineers use two types of apparatus in a wind tunnel. One set of instruments measures the stresses that the moving air induces in the structure being tested. Other equipment shows how the air flows around the object placed in its path. One of four different kinds of apparatus fulfills the latter purpose. The choices are: (1) interferometric instruments such as the Mach-Zehender interferometer, which transforms pressure differences in the air into deflections of the interference fringes purposely produced in light; (2) strioscopic instruments that transform pressure differences in the air into differences of luminosity in the field of observation; (3) systems that produce thin streams of smoke inside the wind tunnel; and (4) slender threads suspended so that they blow in the stream of air. This last system, which is the simplest of all, can be used in the model wind tunnel. If the threads are lightweight and very flexible, they will move with the moving air and show clearly the areas where turbulence occurs.

Illustration 2a shows slender wooden rods with bundles of threads attached to one end. The threads on one rod are 2 cm (about 0.8 in.) long; the threads on the other are 5 cm (about 2 in.) long. A drop of glue fastens the threads to the rods. Thread that is good for this purpose may be obtained by taking warp yarn from a piece of rayon cloth and untwisting it. When the rod is thrust into the wind tunnel, the light, thin threads will indicate directions of airflow.

Illustrations 2b through 2e show the longer threads being used to investigate the movement of air around a cylindrical body that has been placed inside the tunnel so that its axis is perpendicular to the flow of air. In this case the object is a cardboard tube about 6 cm (about 2.4 in.) in diameter. The tube's length is equal to the width of the wind tunnel. The tube is mounted opposite the window, where the camera was positioned to take these photographs.

The rod with its bundle of threads is introduced into the wind tunnel through the access door or through a separate hole drilled in the tunnel wall; it is moved about to investigate air movement in different parts of the tunnel.

Where a smooth flow or streamline movement of the air occurs, the threads take a position and remain more or less motionless until the rod is moved. Where turbulence occurs, on the other hand, the threads show violent agitation as they follow currents of air that move one way and then another.

In these photographs, the threads follow the flow of air around the circular profile of the cylinder. The greatest turbulence is on the side farthest from the source of the air current, where a vortex has developed. Direct observation is far better for this purpose than photographs, because the photographs cannot adequately show the difference between the slight fluttering of the threads in the streamline flow and the violent agitation of the threads in turbulence.

a

b

c

d

e

f

ANALYZING THE FLOW OF AIR AROUND VARIOUS PROFILES—A series of experiments can be performed to investigate the movement of air around objects with different shapes. Illustration 3a shows four different profiles ready for mounting inside the wind tunnel. One is the cylinder that was shown in Illustration 2b; it is made of cardboard and may be a segment of a mailing tube. The semicircular profile may be half of the same kind of cardboard tube, with a flat side glued on. This should be made so that either the rounded or flat side can be placed facing the air current inside the wind tunnel. The other two profiles are teardrop-shaped and resemble cross sections of an aircraft wing but, with somewhat different proportions; one is almost twice as long as it is deep, while the other is about five times as long as it is deep. These objects can be made of wood, cardboard, or some other convenient material; they may be either solid or hollow. It is important that in each case the long dimension be about the same as the inside measurement of the wind tunnel so that the object will extend from the window to the opposite side.

Illustration 3b shows a simple but effective way of making visible the airflow in different parts of the wind tunnel at the same time by mounting many small bundles of threads on the tips of pins. The support is a square piece of thin plywood that fits inside the wind tunnel, parallel to the long axis of the tunnel and facing the window. It is painted black, like the inside of the tunnel. At even intervals of 1 or 2 cm (0.4 or 0.8 in.), strong pins or slender nails are mounted on this support. Bundles of threads 1 to 2 cm long are then fastened to the ends of these pins with glue. Actually, several different boards will be needed to fit around the different profiles.

Illustrations 3c through 3f show the positions the threads attached to the board take when the fan is turned on and air flows around the different profiles. The vortices that develop behind the circular and semicircular profiles are clearly visible. In the case of the semicircular profile, fewer vortices develop when the curved side faces into the stream of air. In the case of the teardrop-shaped profiles, little turbulence develops; there is an almost complete lack of vortices in the area behind the profile where the two streams of air rejoin.

THE ILLUSTRATED SCIENCE DICTIONARY

Silt to Subsonic

KEY TO PRONUNCIATION

The diacritical marks are:

ə b*a*nan*a*, *a*but
ᵊ preceding l, m, n as in batt*le*
ə̇ *e*lectric
ər f*ur*ther
a m*a*t
ā d*ay*
ä cot, f*a*ther
aù n*ow*, *ou*t

e b*e*t
ē b*ea*t
i t*i*p
ī b*i*te
j *j*ob, gem
ŋ si*ng*
ō b*o*ne
ȯ s*aw*, *a*ll
ȯi c*oi*n

th *th*in
t͟h *th*en
ü r*u*le, f*oo*l
u̇ p*u*ll, w*oo*d
ue German h*ü*bsch
u̅e̅ French r*ue*
yü *u*nion
zh vi*si*on

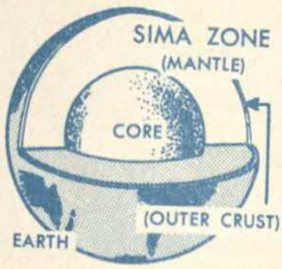
ˈ mark preceding the syllable with strongest stress.
ˌ mark preceding a syllable with secondary stress.

The system of indicating pronunciation in these volumes is used by permission from Webster's *Third New International Dictionary*, copyright 1961 by G. & C. Merriam Co., Publishers of the Merriam-Webster Dictionaries.

silt \'silt\ *n.*

EARTH SCIENCE. Particles of earth materials between 1/256 and 1/16 mm. (.00015 and .0025 inch) in diameter; generally, particles smaller than sand and larger than clay.

SILT is carried by streams and deposited as the streams flow into large bodies of water.



silver \'sil-vər\ *n.*

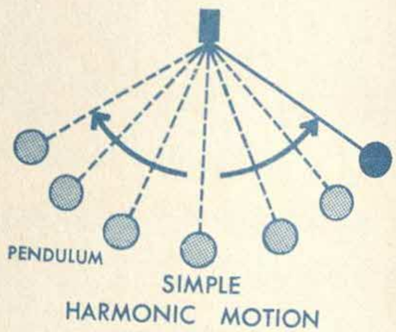
CHEMISTRY. A metallic element that conducts heat and electricity better than any other element and that usually occurs in nature combined with sulfur. Symbol, Ag; atomic number, 47; atomic weight, 107.870.

The ion of SILVER (Ag^+) is the basis for most photographic films because it is sensitive to light.

sima zone \'sī-mə 'zōn\

EARTH SCIENCE. A layer or shell of dense, igneous rock, theorized to act as support for the less-dense sial zone and believed to contain silicon and magnesium, symbols Si and Ma, from which it is named.

The SIMA ZONE is considered to have a specific gravity of about 3.3.



simple harmonic motion \'sim-pəl här-'män-ik 'mō-shən\

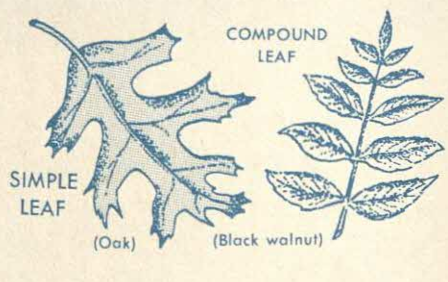
PHYSICS. A regular back-and-forth or up-and-down motion. The force causing the motion always acts toward the rest position of the object, and the amount of force is proportional to the distance of the object from its rest point.

SIMPLE HARMONIC MOTION is approximated by a swinging pendulum.

simple leaf \'sim-pəl 'lēf\

BOTANY. A leaf composed of a single blade; see *compound leaf.*

An oak leaf is a SIMPLE LEAF.



simple sugar \'sim-pəl 'shu̇g-ər\

CHEMISTRY. Any one of several sweet-tasting compounds having the chemical formula $C_6H_{12}O_6$. They are often called monosaccharides, as distinguished from polysaccharides that have more complex molecules; see *monosaccharide, disaccharide* and *polysaccharide.*

In digestion, a molecule of table sugar (sucrose) is changed into a molecule of glucose and a molecule of fructose, each of which is a SIMPLE SUGAR.

simulation \\ˌsim-yə-ˈlā-shən\\ *n.*

ZOOLOGY. A pattern of behavior in which an organism imitates another species or pretends to be injured; mimicking or feigning.

The opossum's response to danger is a SIMULATION *of death, referred to as "playing possum."*

sine \\ˈsīn\\ *n.*

MATHEMATICS. In trigonometry, a function of an angle or of the real number that is the radian measure of that angle. Expressed as a function of an acute angle of a right triangle, it is equal to the ratio of the side opposite the given angle to the hypotenuse. Expressed as a function of an angle in standard position in a coordinate plane, it is the ratio of the ordinate of a point on the terminal side of the angle to the distance from the origin to the point; *abbr.* sin.

The SINE *of an angle of 30 degrees and the cosine of an angle of 60 degrees are both ½.*

SINK HOLE

sine wave \\ˈsīn ˈwāv\\

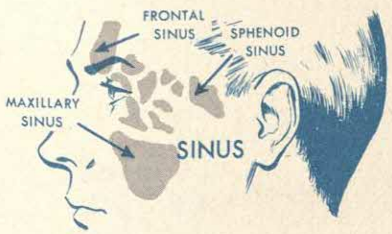
PHYSICS. A wave that, when drawn on a graph with displacement plotted against time or distance, follows a trigonometric sine curve; any harmonic wave produced by elastic matter vibrating with harmonic motion.

The amplitude of displacement at any point on a SINE WAVE *is proportional to the sine of the phase angle of the displacement.*

single replacement \\ˈsiŋ-gəl ri-ˈplā-smənt\\

CHEMISTRY. A type of chemical reaction in which an atom of one element takes the place of an atom or ion of another element in a compound.

A SINGLE REPLACEMENT *reaction between iron and copper chloride gives copper and iron chloride.*



SINUS

sinkhole \\ˈsiŋk-ˌhōl\\ *n.*

EARTH SCIENCE. A hole or depression in the ground, created by the dissolving action of water on such soluble rocks as limestone and gypsum; also called a sink.

A SINKHOLE *usually connects with an underground drainage system.*

sinus \\ˈsī-nəs\\ *n.*

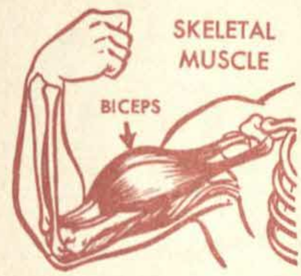
ANATOMY. A cavity or hollow space within a structure of the body; also, an enlarged or dilated channel, as those of veins in the cranium; especially, the hollows in the bones of the skull.

Inflammation of a SINUS *in the facial bones may result from a cold.*

siphon \'sī-fən\ *n.*

1. PHYSICS. A bent tube, generally shaped like an upside-down U, with one arm longer than the other. It uses air pressure and gravity to move fluid over a barrier to a lower level. 2. ZOOLOGY. In certain mollusks, a tubular-shaped organ that carries water to the gills or that expels liquid from the gill chamber.

A SIPHON *will not work in a vacuum.*



skeletal muscle \'skel-ət-ᵊl 'məs-əl\

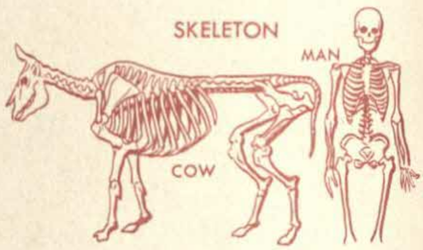
ANATOMY. A voluntary muscle, connected to the bones, that moves some part of the body.

A SKELETAL MUSCLE *is usually opposed by another muscle called an antagonistic muscle.*

skeleton \'skel-ət-ᵊn\ *n.*

1. ANATOMY and ZOOLOGY. The rigid framework of an animal's body, usually jointed to allow movement. In vertebrate animals, it is an internal framework of bone, of cartilage or of bone and cartilage. 2. The supporting framework of a structure.

Many invertebrate animals, such as insects and crustaceans, have an external SKELETON, *or exoskeleton.*



skull \'skəl\ *n.*

ANATOMY and ZOOLOGY. That part of the skeleton that makes up the bony part of the head of vertebrate animals.

In man, the SKULL *is made up of 8 cranial bones and 14 facial bones.*

slant height \'slant 'hīt\

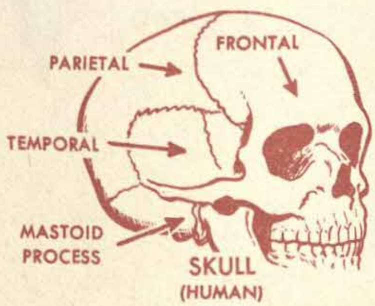
MATHEMATICS. The altitude of any lateral face of a regular pyramid or frustum of a regular pyramid; of a right circular cone, the length of any one of its elements.

The lateral area of a regular pyramid is equal to one half the product of the perimeter of its base and its SLANT HEIGHT.

slate \'slāt\ *n.*

EARTH SCIENCE. A metamorphic rock that is formed by metamorphism of shale.

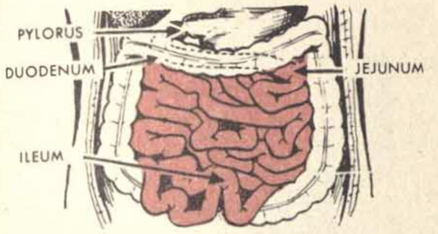
Although SLATE *splits easily into thin layers, it is difficult to split in any other direction.*



sleet \'slēt\ *n.*

EARTH SCIENCE. Rain that has been frozen, or partly frozen, into ice particles; also, a mixture of rain and snow.

SLEET *can create hazardous driving conditions by making roads slippery and by reducing visibility.*



SMALL INTESTINE

small intestine \'smȯl in-'tes-tən\
ANATOMY. The coiled, tubular part of the digestive tract in the abdomen. It consists of the duodenum, jejunum and ilium, and extends from the pylorus of the stomach to the large intestine.

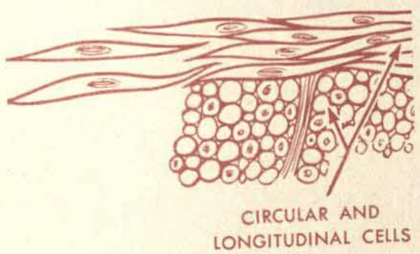
Digestion and absorption of food occur in the SMALL INTESTINE.

smelting \'smelt-iŋ\ *n.*
ENGINEERING. A process by which certain metals are obtained from their ores. It usually involves heating an ore to a high temperature in the presence of a reducing agent.

SMELTING *of iron ore is often done by heating a mixture of ore and coke (carbon) in a blast furnace.*

smog \'smȯg\ *n.*
EARTH SCIENCE. A combination of smoke and fog, most common in industrial areas near rivers, lakes or oceans.

The fumes and smoke particles in SMOG *may cause respiratory irritation.*



SMOOTH MUSCLES

smoke \'smōk\ *n.*
CHEMISTRY and ENGINEERING. A suspension of very small, solid particles in a gas. Smoke is frequently produced by the incomplete burning of coal, wood or oil; see *smog.*

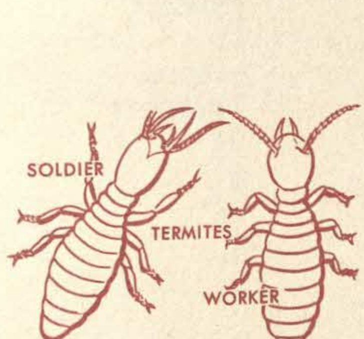
Most large industrial cities have established laws to control and limit the discharge of SMOKE.

smooth muscles \'smüt͟h 'məs-əlz\
ANATOMY. The involuntary muscles characterized by smooth or unstriated fibers, as in the walls of a digestive organ.

SMOOTH MUSCLES *function without conscious, or voluntary, control.*

snow \'snō\ *n.*
EARTH SCIENCE. A form of precipitation in which water drops are frozen into ice crystals having a variety of six-sided shapes. Snow may fall as separate crystals or as crystals clumped together.

A blanket of SNOW *may protect vegetation during long periods of below-freezing temperatures.*



SOCIAL ANIMAL

social animal \'sō-shəl 'an-ə-məl\
ZOOLOGY. An animal that characteristically associates with a group of animals of the same species for most, or all, of its life

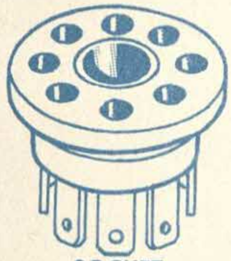
cycle and that performs a role that is part of the total activities carried out by the organization or society.

A SOCIAL ANIMAL *may be part of a family group, as is the soldier termite, or part of a grouping of families, as is the sentinel gopher in a gopher colony.*

socket \'säk-ət\ *n.*

1. ANATOMY. A hollow or cavity in the body that contains another part of the body, as the socket of the eye. 2. ENGINEERING. A hollow device used to hold or support a piece of electrical equipment, such as an electric light or electric plug, and to connect the electrical equipment to a source of current.

The human eye is protected by a bony SOCKET.



SOCKET

soda \'sōd-ə\ *n.*

CHEMISTRY. Any one of several compounds of sodium; frequently, sodium carbonate.

Baking SODA *(sodium hydrogen carbonate) reacts with acid to produce carbon dioxide.*

soft water \'sȯft 'wȯt-ər\

CHEMISTRY. Water that is relatively free from compounds of calcium and magnesium.

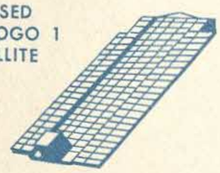
Soap forms suds easily in SOFT WATER.

soil \'sȯi(ə)l\ *n.*

BOTANY and EARTH SCIENCE. That part of the earth's surface containing a mixture of rock particles, minerals and organic matter that will support plant growth; generally, all loose, weathered material above the bedrock.

The kind of SOIL *in a region depends largely on the geological history, the climate and the life forms of the region.*

AS USED ON OGO 1 SATELLITE



SOLAR BATTERY

sol \'säl\ *n.*

BIOLOGY and CHEMISTRY. A solution of colloidal particles in a clear liquid, distinguished from true solutions by the fact that a beam of light is visible in it; see *colloid* and *Tyndall effect.*

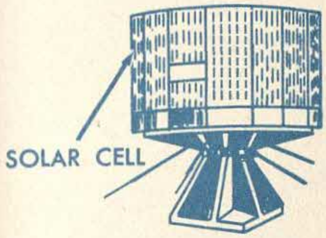
The cytoplasm in a muscle cell is a SOL *that changes to a gel when the muscle contracts.*

solar battery \'sō-lər 'bat-ə-rē\

ASTRONAUTICS. A battery operated by the sun's rays. Silicon cells in the battery produce electricity when they are struck by the sun's rays. A solar battery can produce enough electricity

while it is in sunlight to operate electrical equipment and to recharge conventional batteries.

A SOLAR BATTERY *may be used to provide power for radio equipment carried by an artificial satellite.*



solar cell \'sō-lər 'sel\

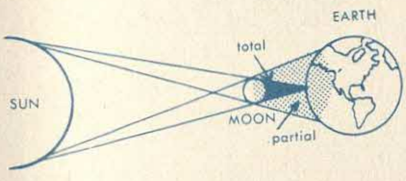
ENGINEERING. A device that uses radiant energy from the sun to produce an electric current.

A silicon SOLAR CELL *in bright sunlight produces about ¼ volt.*

solar day \'sō-lər 'dā\

ASTRONOMY. Either of two periods of time, one based on two successive transits of the apparent sun across an observer's meridian (sundial time) and the other based on two successive transits of a mean sun across an observer's meridian (clock time); see *mean solar time* and *sidereal day.*

Because the earth's angular rate of revolution around the sun is not constant, the SOLAR DAY *based on the apparent sun varies in length and therefore is not used for most purposes.*



SOLAR ECLIPSE

solar eclipse \'sō-lər i-'klips\

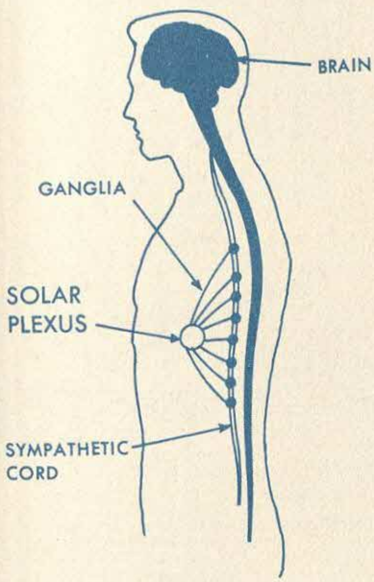
ASTRONOMY. An eclipse that occurs when the moon passes between the sun and the earth, causing the moon's shadow to fall on the earth's surface.

Until the invention of the coronagraph, astronomers could study the sun's corona only during a SOLAR ECLIPSE.

solar energy \'sō-lər 'en-ər-jē\

ASTRONOMY. The radiant energy of the sun.

SOLAR ENERGY *can produce electricity for orbiting artificial satellites.*



solar plexus \'sō-lər 'plek-səs\

ANATOMY. The center of the sympathetic nervous system. It is located in the upper abdomen behind the stomach and in front of the aorta.

A blow in the SOLAR PLEXUS *may cause the heart to beat so slowly that it cannot pump enough blood to the brain to maintain consciousness.*

solar radiation \'sō-lər ˌrād-ē-'ā-shən\

ASTRONOMY. Energy released from the sun. One type of solar radiation is electromagnetic radiation of all wavelengths. Another type is corpuscular radiation, which is made up of high-speed particles, mostly electrons and protons.

When the number of sunspots increases, SOLAR RADIATION *becomes greater.*

solar system \'sō-lər 'sis-təm\

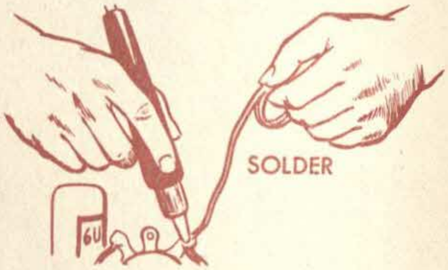
ASTRONOMY. The sun, the nine planets, the satellites of the planets, the asteroids and the meteors and comets; generally, all celestial bodies within the orbit of Pluto.

One theory of the origin of the SOLAR SYSTEM *states that the sun and other bodies arose from a cloud of whirling gas and dust.*

solar time \'sō-lər 'tīm\

ASTRONOMY. Time based on either an apparent sun or a mean sun; see *solar day* and *mean solar time.*

SOLAR TIME *is based on the interval of time required for the earth to make one rotation with respect to the sun.*



SOLDER

solder \'säd-ər\ *n.*

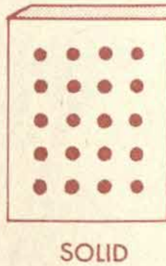
CHEMISTRY and ENGINEERING. An alloy of low melting point, used to join two pieces of metal or wire; usually, an alloy of lead and tin.

Acid-core SOLDER *is used on sheet metal, but rosin-core solder is used on copper wire.*

solid \'säl-əd\ *n.*

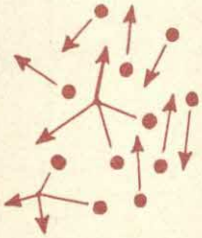
1. PHYSICS. Any substance with definite shape and a nearly-constant volume, that, unlike liquids and gases, resists forces tending to change its shape. 2. MATHEMATICS. A three-dimensional figure.

The molecules in a SOLID *are rather rigidly fixed in place, while liquid and gas molecules are relatively free to move.*



SOLID

LIQUID



GAS

solid geometry \'säl-əd jē-'äm-ə-trē\

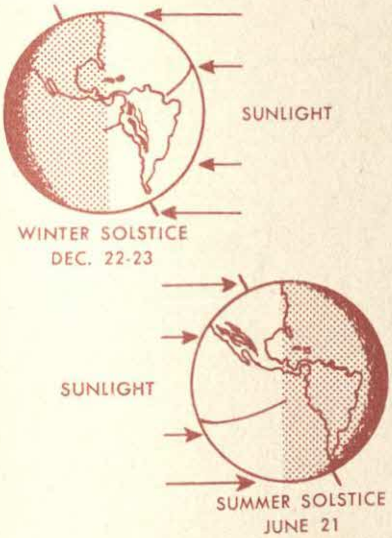
MATHEMATICS. That part of geometry dealing with the study of figures in three-dimensional space, and the logical investigation and study of the properties, characteristics and relationships of plane and solid figures in three-dimensional space.

The study of SOLID GEOMETRY *is of particular benefit to students in most branches of engineering.*

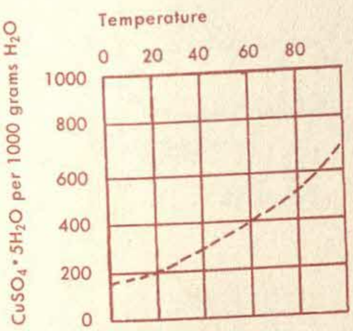
solid propellant \'säl-əd prə-'pel-ənt\

ASTRONAUTICS. A rocket propellant in solid form. It consists of one or more chemical compounds and usually contains both fuel and oxidizer; see *liquid propellant.*

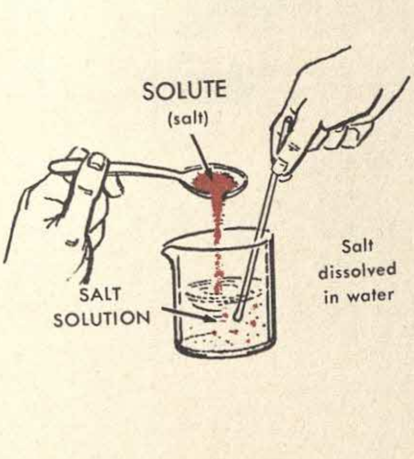
Small rockets loaded with SOLID PROPELLANT *are often used as distress signals.*



SOLSTICE



SOLUBILITY
OF COPPER SULPHATE



solid solution \'säl-əd sə-'lü-shən\

CHEMISTRY. A uniform mixture of two or more substances in the solid state. The relative amounts of its components are variable, as contrasted with a compound, in which the relative amounts are fixed. Frequently, a solid solution is an alloy, such as brass or stainless steel.

Steel is a SOLID SOLUTION *of iron, carbon and other metals, such as manganese and nickel.*

solid state physics \'säl-əd 'stāt 'fiz-iks\

PHYSICS. The branch of science that deals with the physical properties and structures of solids.

Research in SOLID STATE PHYSICS *has yielded information used in the development of the transistor.*

solstice \'säl-stəs\ *n.*

ASTRONOMY. One of two times each year (the first day of summer and the first day of winter) when the sun appears to reach its maximum distance from the celestial equator; also, either of the two points on the ecliptic that is the maximum distance from the celestial equator.

The summer SOLSTICE *occurs about June 22, while the winter solstice occurs about December 22.*

solubility \ˌsäl-yə-'bil-ət-ē\ *n.*

CHEMISTRY and PHYSICS. A measure of the amount of a given substance that will dissolve in another substance; also, the greatest weight of a given substance that will dissolve at a specified temperature in a given weight of water.

While the SOLUBILITY *of table salt in water is about 36 grams per 100 grams of water at 20° C., the solubility increases as the temperature increases.*

solute \'säl-ˌyüt\ *n.*

CHEMISTRY and PHYSICS. A substance that is dissolved in another substance, or solvent, to form a solution, as salt or sugar that is dissolved in water; also, any one of the components of a solution that is present in a smaller amount than the substance (solvent) making up the greatest part.

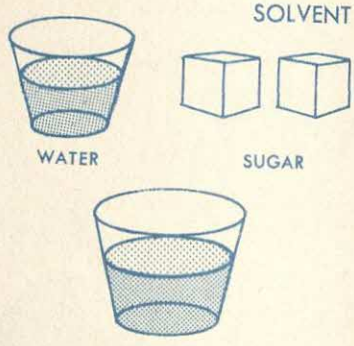
Although seawater contains calcium and magnesium compounds, ordinary table salt (sodium chloride) is the principal SOLUTE *in seawater.*

solution \sə-'lü-shən\ *n.*

1. CHEMISTRY and PHYSICS. A uniform mixture of two substances that cannot be separated by settling, filtering or other mechani-

cal means; a mixture of molecules or smaller particles. 2. MATHEMATICS. The process, or step-by-step procedure, followed in determining a result that is necessary and sufficient to meet the conditions of a problem.

A SOLUTION *of alcohol and water may be made in any proportion, while the proportions of a sugar-and-water solution are limited by the solubility of sugar.*

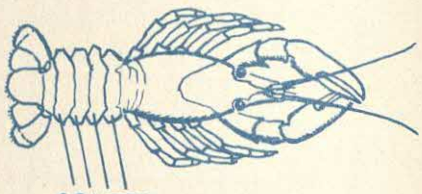


solvation \säl-'vā-shən\ *n.*
CHEMISTRY. The attraction and bonding between dissolved ions and the molecules of the substance in which they are dissolved.

When acids dissolve in water, any free hydrogen ions that are formed undergo SOLVATION *with water molecules.*

solvent \'säl-vənt\ *n.*
CHEMISTRY and PHYSICS. The substance in which another substance dissolves to form a solution; frequently, a liquid, such as water or alcohol, that will dissolve a large number of substances; also, the component of a solution that makes up more of the solution than any other component.

Water is a good SOLVENT *for many crystalline substances, such as salt and sugar, but it is a poor solvent for covalent substances, such as rubber and sulfur.*

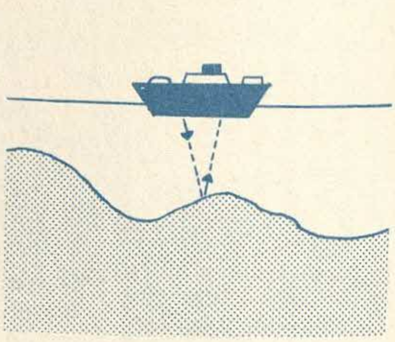


somatic \sō-'mat-ik\ *adj.*
BIOLOGY. Referring to all portions of the body of an organism except the germ cells involved in sexual reproduction.

A SOMATIC *mutation in a parent plant or animal cannot be passed on to its offspring unless that plant or animal reproduces asexually.*

somatoplasm \'sō-mət-ə-ˌplaz-əm\ *n.*
BIOLOGY. All the body cells, except the germ cells (germ plasm), that make up the bulk of an organism.

The SOMATOPLASM *of an organism eventually dies, but the germ plasm may be passed on through sexual reproduction.*



somites \'sō-ˌmīts\ *n.*
1. ZOOLOGY. A series of segments extending along an animal's body, as in insects and earthworms. 2. ANATOMY. The incompletely-formed segments in the bodies of developing embryos.

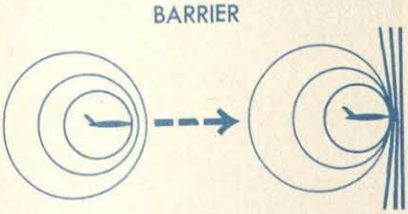
SOMITES *can be observed on the abdomen of a crayfish.*

sonar \'sō-ˌnär\ *n.*
ENGINEERING. A device used aboard ships to map the ocean

floor or to locate submerged objects by sending out an underwater sound wave and recording the time required for the sound to echo back to the ship. The name is a contraction of "*so*und *na*vigation *r*anging."

Commercial fishermen use SONAR *to locate schools of fish.*

SONIC BARRIER



As aircraft speed approaches speed of sound, air pressure waves pile up to form a wall.

sonic barrier \'sän-ik 'bar-ē-ər\
AERONAUTICS and ASTRONAUTICS. The increasing air resistance that is encountered by an aircraft as it approaches the speed of sound.

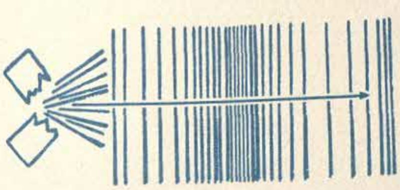
Supersonic airplanes are designed with thin wings and slim bodies to aid penetration of the SONIC BARRIER.

sonic boom \'sän-ik 'büm\
AERONAUTICS. The sound resulting from a shock wave caused by an aircraft or other object moving at or above the speed of sound.

The sound wave from a SONIC BOOM *produced at low altitudes will sometimes shatter window glass.*

sorter \'sȯrt-ər\ *n.*
ENGINEERING. That part of a punched card data processing system that sorts key punch cards according to the location of the holes in them.

One type of SORTER *in common use will process a maximum of 650 cards per minute.*



Noise causes compression of air

SOUND

sound \'saůnd\
1. PHYSICS (N.). A series of pressure changes produced by a vibrating object and transmitted by a medium such as air or water. 2. PHYSIOLOGY (N.). The effect of vibrations in the air, or other media, on the organs of hearing. 3. EARTH SCIENCE (N.). A narrow strip or passage of water connecting two large bodies of water or lying between an island and the mainland; also, a large strait. (V.). To determine water depth with a weighted line or with sonar equipment.

The speed of SOUND *is greater in liquids and solids than in gases.*

sound barrier \'saůnd 'bar-ē-ər\
AERONAUTICS and ASTRONAUTICS. A popular term for sonic barrier. See *sonic barrier.*

sounding rocket \'saůn-diŋ 'räk-ət\
ASTRONAUTICS. A kind of rocket designed to study the upper

atmosphere. Information is gained by bouncing radio waves off particles in the atmosphere.

A SOUNDING ROCKET *is used to determine the density of the ionosphere.*

SOUTHERN HEMISPHERE

Southern Hemisphere \'səth-ərn 'hem-ə-ˌsfi(ə)r\

EARTH SCIENCE. That half of the earth that lies south of the equator.

Australia is in the SOUTHERN HEMISPHERE.

southern lights \'səth-ərn 'līts\

ASTRONOMY. The popular term for aurora australis. See *aurora australis.*

SOUTHERN LIGHTS

south magnetic pole \'sau̇th 'mag-net-ik 'pōl\

EARTH SCIENCE. The place on the earth's surface toward which the south-seeking pole of a compass points, located 72 degrees south latitude and 155 degrees east longitude.

The earth's SOUTH MAGNETIC POLE *is located on the continent of Antarctica.*

South Pole \'sau̇th 'pōl\

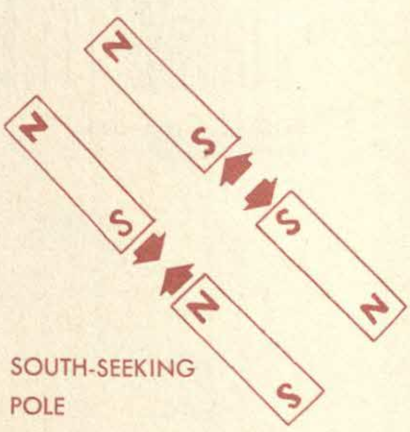
EARTH SCIENCE. The southernmost point on the earth. It is one of two points through which the imaginary line forming the earth's axis passes and about which the earth rotates. It is also called the geographical south pole to distinguish it from the south magnetic pole.

The discovery of coal in Antarctica supports the belief that the SOUTH POLE *once had a tropical climate.*

south-seeking pole \'sau̇th-'sēk-iŋ 'pōl\

EARTH SCIENCE and PHYSICS. The end of a magnetized bar, such as a compass needle, that points south; often designated as S-pole.

If a SOUTH-SEEKING POLE *is placed in the magnetic field of another south-seeking pole, the two poles will repel each other.*



SOUTH-SEEKING POLE

space charge \'spās 'chärj\

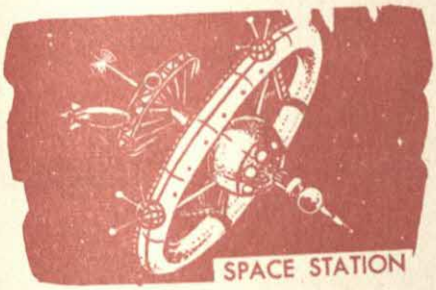
PHYSICS. Electrons given off when a filament is heated. When a filament emits electrons, it becomes positively charged, holding some electrons in the near vicinity of the filament.

A diode vacuum tube contains a positively-charged anode, which attracts the SPACE CHARGE.

space probe \'spās 'prōb\

ASTRONAUTICS. A spacecraft sent to the vicinity of another planet or natural satellite or placed in orbit between planets. It is equipped with detection equipment and a radio transmitter that relays to earth information about conditions in space.

Mariner II, a United States SPACE PROBE, *passed within 21,000 miles of Venus on December 14, 1962.*



space station \'spās 'stā-shən\

ASTRONAUTICS. A proposed manned artificial satellite that will serve as a takeoff point for long-range space trips, as a relay station for broadcasting and, generally, as a research laboratory; also called a space platform.

According to one plan, the sections of a SPACE STATION *will be placed in orbit and assembled there by a team of astronauts.*

space-time continuum \'spās-'tīm kən-'tin-yə-wəm\

PHYSICS. A concept of space, involving four dimensions to locate an event. The dimensions correspond to length, width, height and time.

If two moving objects arrive at exactly the same position in the SPACE-TIME CONTINUUM, *they will collide.*

spawn \'spȯn\ *n.*

ZOOLOGY. A mass of eggs deposited by fish, amphibians or other aquatic, oviparous animals.

The SPAWN *of the lobster remains attached to the female for approximately ten months.*

specialization \ˌspesh-(ə-)lə-'zā-shən\ *n.*

BIOLOGY. The adaptation of a part of an organism to a particular function; also, the adaptation of an organism to its environment.

The evolutionary development of the seal's forelimbs into flippers illustrates SPECIALIZATION.

speciation \ˌspē-s(h)ē-'ā-shən\ *n.*

BIOLOGY. The evolutionary formation of two or more species arising from one common ancestral species, as when an interbreeding population is divided and prevented from remixing for many generations by some sort of natural or artificial barrier.

SPECIATION *apparently occurred in Colorado after one population of squirrels was divided by the formation of the Grand Canyon, resulting in the evolution of at least two squirrel species.*

species \'spē-(ˌ)shēz\ *n.*

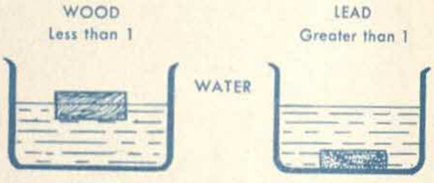
BIOLOGY. A group of plants or animals so similar in structure and heredity traits that their various forms will normally interbreed for successive generations. Species is the classification unit below genus.

Modern man is of the SPECIES *sapiens and is the only known living member of the genus* Homo.

specific gravity \spi-'sif-ik 'grav-ət-ē\

CHEMISTRY and PHYSICS. A measure of the relative heaviness or lightness of a substance. It is the ratio of the weight of an object to the weight of an equal volume of water.

Substances with a SPECIFIC GRAVITY *of less than one will float in water and are said to be lighter than water.*



SPECIFIC GRAVITY

specific heat \spi-'sif-ik 'hēt\

PHYSICS. The amount of heat, expressed as the number of calories, that, when absorbed by one gram of a substance, will raise the temperature of the substance one degree Centigrade.

The SPECIFIC HEAT *of water is one, while for all other common solids and liquids it is less than one.*

specific impulse \spi-'sif-ik 'im-ˌpəls\

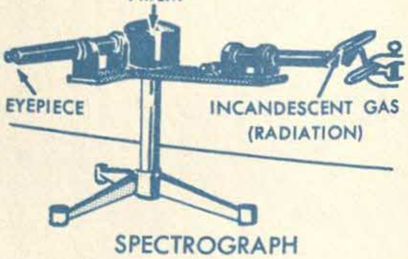
ASTRONAUTICS. A measure of rocket-engine efficiency that relates the thrust produced to the pounds of propellant used and the burning time required.

In designing a rocket, SPECIFIC IMPULSE *and mass ratio are important considerations.*

spectral classes of stars \'spek-trəl 'klas-əz əv 'stärz\

ASTRONOMY. A grouping, or ranking, of stars according to their spectra. The classification may also be considered one of temperature, since there is a direct relationship between a star's temperature and its spectrum. Spectral classes are usually designated by the letters O, B, A, F, G, K and M; see *spectrum-luminosity diagram.*

Among the SPECTRAL CLASSES OF STARS, *O stars are hottest (50,000° K.) and M stars are coolest (3,200° K.).*

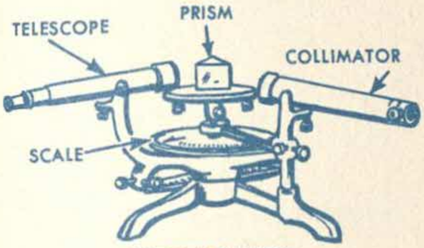


SPECTROGRAPH

spectrograph \'spek-trə-ˌgraf\ *n.*

PHYSICS. A device used in some types of chemical analysis for producing and recording spectra. It produces spectra by a transparent prism through which radiation passes or by a diffraction grating.

A SPECTROGRAPH *attached to a telescope provides data on the chemical composition of stars.*



SPECTROMETER

spectrometer \spek-'träm-ət-ər\ *n.*

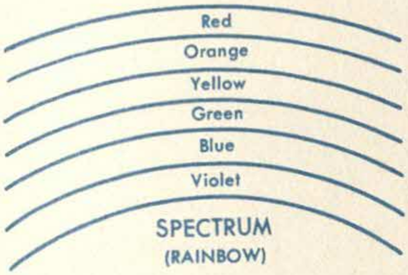
PHYSICS. An instrument used to measure and study spectra and to measure the wavelengths of light from various sources.

A SPECTROMETER *may be used to analyze ultraviolet and infrared radiation, as well as visible light.*

spectroscopy \spek-'träs-kə-pē\ *n.*

PHYSICS. That branch of science concerned with the study of the spectrum, including photographic observation of the spectrum and the measurement and determination of wavelengths.

In SPECTROSCOPY, *one of the instruments widely used is a diffraction-grating spectroscope.*



SPECTRUM
(RAINBOW)

spectrum \'spek-trəm\ *n.*

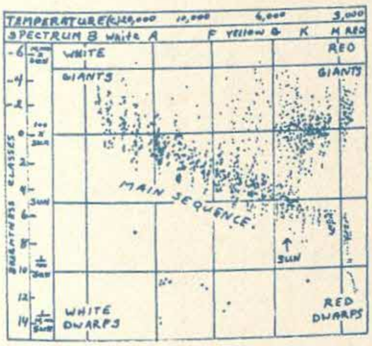
PHYSICS. A visual image, as in a rainbow, of the colors that make up white light; also, a separation by wavelength of energy, such as light or X rays, or a separation of a stream of subatomic particles, such as alpha or beta radiation, by the different energies of the particles; also, an analysis of all electromagnetic radiation into its component parts; see chart, page 624.

A SPECTRUM *may be produced by refraction or diffraction of light.*

spectrum-luminosity diagram \'spek-trəm ˌlü-mə-'näs-ət-ē 'dī-ə-ˌgram\

ASTRONOMY. A graph or diagram obtained by plotting absolute magnitudes of stars against their spectral classes. The diagram is useful in classifying stars and in explaining the possible evolution of stars; see *main-sequence stars, dwarf star* and *giant star.*

A SPECTRUM-LUMINOSITY DIAGRAM *is also called a Hertzsprung-Russell diagram.*



SPECTRUM-LUMINOSITY DIAGRAM

speed \'spēd\ *n.*

PHYSICS. The rate at which a moving object travels; the distance traveled in a given unit of time, such as a second or minute.

If a car travels 80 miles in 2 hours, its average SPEED *is 40 miles per hour.*

speed of light \'spēd əv 'līt\

PHYSICS. The rate at which light moves, used as a constant in certain theories. In a vacuum, the rate is approximately 186,000 miles per second (3×10^8 meters per second).

The SPEED OF LIGHT *in water or glass is less than its speed in a vacuum.*

speed of sound \'spēd əv 'saůnd\
PHYSICS. The rate at which sound travels through a medium. The rate varies with the kind of medium and with the temperature and density of the medium. It is about 741 miles per hour, or 1,087 feet per second, through air at 0° C. and at normal atmospheric pressure.

Airplanes approaching the SPEED OF SOUND *go through a temporary condition of vibration and instability.*

sperm \'spərm\ *n.*
BIOLOGY. The male reproductive cell that fuses with an egg, or female reproductive cell, to form a zygote.

A SPERM *can usually be distinguished from the female gamete, or ovum, by special structures, such as cilia or a flagellum.*

spermatogenesis \ˌspər-mət-ə-'jen-ə-səs\ *n.*
BIOLOGY. The formation of sperm in the male germ tissue.

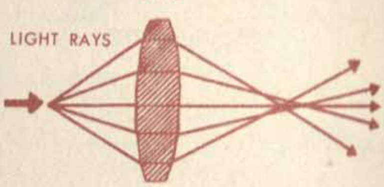
The end products of SPERMATOGENESIS *are motile sperm cells.*

spermatophytes \(ˌ)spər-'mat-ə-ˌfīts\ *n.*
BOTANY. Plants that produce pollen tubes and seeds and that have true roots, stems and leaves.

Two major groups of SPERMATOPHYTES *are the cone-bearing plants and the flowering plants.*

spherical aberration \'sfir-i-kəl ˌab-ə-'rā-shən\
PHYSICS. The blurring of an image formed by a lens or curved mirror due to the fact that light reflected or refracted at the edge of the lens or mirror is directed to a different focus from that of light from the center of the lens or mirror.

The diaphragm of a camera eliminates much SPHERICAL ABERRATION *by cutting out light rays that pass through the edge of the lens.*



SPHERICAL ABERRATION

sphincter muscle \'sfiŋ(k)-tər 'məs-əl\
ANATOMY and ZOOLOGY. In man and most animals, a ring of muscle fiber that closes a passage or opening when constricted.

The cardiac SPHINCTER MUSCLE *controls the opening from the esophagus to the stomach.*

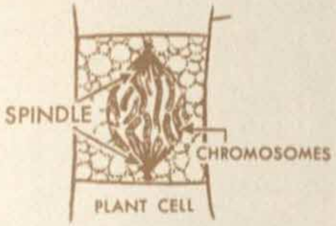
sphygmomanometer \ˌsfig-(ˌ)mō-mə-'näm-ət-ər\ *n.*
MEDICINE. An instrument used to measure arterial blood pressure. It consists of an inflatable pressure cuff attached to a pressure gauge.

A SPHYGMOMANOMETER *measures both systolic and diastolic blood pressure.*

SPHYGMOMANOMETER

spinal column \'spīn-ᵊl 'käl-əm\
ANATOMY. That part of the skeleton of vertebrate animals that is flexible and made up of a series of vertebrae that are connected by fibrous disks and ligaments.

The SPINAL COLUMN *is the main axis of the skeleton.*

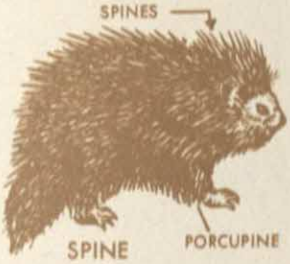


spinal cord \'spīn-ᵊl 'kȯ(ə)rd\
ANATOMY and PHYSIOLOGY. In vertebrate animals, a thick, rope-like structure within the spinal column. The cord functions as the main conductor of nerve impulses to and from the brain and as a center for many independent reflex actions.

The SPINAL CORD *contains nerve cell bodies, many nerve fibers and supporting cells, or neuroglia.*

spindle \'spin-dᵊl\ *n.*
BIOLOGY. A framework of colorless fibers, narrowed in both directions from a wide middle section, that appears in cells undergoing mitosis.

During anaphase, fibers of the SPINDLE *attached to the chromosome pairs appear to pull the chromosomes to opposite poles of the dividing cell.*

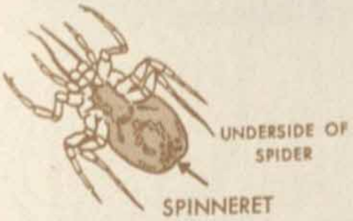


spine \'spīn\ *n.*
BIOLOGY. A stiff, sharp-pointed external growth, such as the pointed ray of a catfish or one of the projections from the stem of a thistle.

A porcupine quill, or SPINE, *can be raised by muscle action but cannot be thrown.*

spinneret \ˌspin-ə-'ret\ *n.*
1. ZOOLOGY. A tubular, silk-expelling organ on the abdomen of a spider or near the mouth of a caterpillar. 2. ENGINEERING. A device containing many small openings through which liquid is forced in manufacturing synthetic fibers.

The end of a typical SPINNERET *contains a large number of small pores.*



spiracle \'spī-ri-kəl\ *n.*
ZOOLOGY. An external opening of the breathing system, especially a pore from a trachea along the abdomen and thorax of an insect.

In certain species of insects, the SPIRACLE *is covered by a framework of hairs that filter dust.*

spiral galaxy \\'spī-rəl 'gal-ək-sē\\

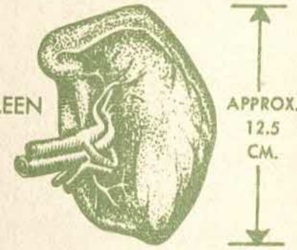
ASTRONOMY. A huge, rotating star system containing a prominent nucleus of stars centered in a disk of stars and gaseous dust in the form of spiral arms. It is sometimes called a spiral nebula and usually contains more than a billion stars.

The earth is located near the edge of a SPIRAL GALAXY *known as the Milky Way.*

spit \\'spit\\ *n.*

EARTH SCIENCE. A type of coastal sandbar that extends outward from a shore into open water.

A SPIT *is formed by sand deposits from ocean currents that flow nearly parallel to the shoreline.*



spleen \\'splēn\\ *n.*

ANATOMY and ZOOLOGY. A large, glandular organ found in most vertebrate animals. In man, it is located below the left side of the diaphragm and behind the upper part, or fundus, of the stomach and functions as a part of the lymphatic system.

The SPLEEN *forms blood cells, releases hemoglobin by breaking down old red corpuscles, filters harmful materials from the blood, stores blood and provides the blood with antibodies.*

spontaneous combustion \\spän-'tā-nē-əs kəm-'bəs-chən\\

CHEMISTRY. Burning that starts from heat formed within an object by slow oxidation of a substance in the object.

SPONTANEOUS COMBUSTION *of oily rags is caused by heat produced from the slow action of air on oil.*

spontaneous generation \\spän-'tā-nē-əs ˌjen-ə-'rā-shən\\

BIOLOGY. A theory, now disproved, that living organisms found in decayed or dead organic matter were produced by and from such matter; also called abiogenesis.

SPONTANEOUS GENERATION *was once considered the explanation for the appearance of maggots in meat.*

sporangium \\spə-'ran-jē-əm\\ *n.*

BIOLOGY. A walled chamber or case in which asexual spores are formed in algae, fungi, mosses, ferns and certain protozoa.

A SPORANGIUM *of bread mold fungus darkens and then breaks open as the spores in it mature.*

spore \\'spō(ə)r\\ *n.*

BIOLOGY. A reproductive body that develops into a separate organism without fertilization. In plants, a spore is usually pro-

duced by meiosis and gives rise to a haploid gametophyte. In protozoa and bacteria, it may be formed by division, by enclosure of the protoplasm in a resistant covering or by both division and enclosure.

A pollen grain is a SPORE *that germinates, producing a male plant that, in turn, produces sperm cells.*

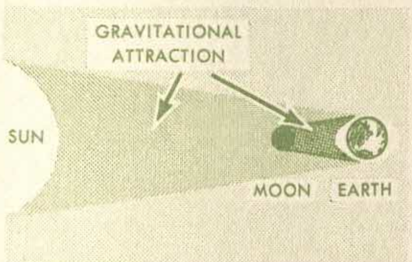
SPOROPHYTE

sporophyte \ˈspōr-ə-ˌfīt\ *n.*
BOTANY. That stage in the life cycle of a spermatophyte, moss or fern that develops from a fertilized egg and in which spores are produced by meiosis; see *gametophyte.*

The SPOROPHYTE *has evolved to become the dominant stage among land plants.*

sport \ˈspō(ə)rt\ *n.*
BIOLOGY. Any organism that differs strikingly from its parents, usually caused by mutation.

A tailless dog born of parents with tails is a SPORT.



SPRING TIDE

spring tide \ˈspriŋ ˈtīd\
EARTH SCIENCE. The tide that rises highest above, and falls lowest below, mean sea level. It occurs twice a month, near the times of new and full moon.

During a SPRING TIDE, *the sun, earth and moon are approximately in a straight line.*

squall line \ˈskwȯl ˈlīn\
EARTH SCIENCE. A long line of towering cumulus and cumulonimbus clouds that occurs in front of a fast-moving cold front. A squall line is frequently the source of violent thunderstorms and, occasionally, of tornadoes.

A SQUALL LINE *looks like a wall of black clouds rolling forward and boiling upward.*

square \ˈskwa(ə)r\ *n.*
MATHEMATICS. In geometry, a four-sided figure whose sides are equal and whose angles are equal; also, in algebra and arithmetic, the result of a quantity multiplied by itself.

The area of a SQUARE *is equal to its length multiplied by its width.*

square number \ˈskwa(ə)r ˈnəm-bər\
MATHEMATICS. A number that is the square of a whole number, or an integer.

The number 25 is a SQUARE NUMBER, *since it is obtained by squaring 5.*

square root \'skwa(ə)r 'rüt\

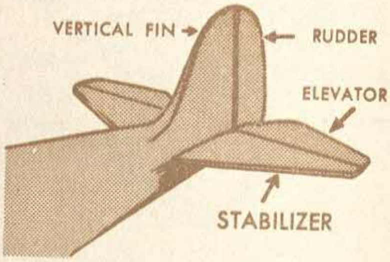
MATHEMATICS. A number or quantity that, if multiplied by itself, produces a given number or quantity.

The positive SQUARE ROOT *of 9 is 3.*

square wave \'skwa(ə)r 'wāv\

PHYSICS. A wave pattern characterized by regular but abrupt, rather than gradual, transitions from one value to another.

When a SQUARE WAVE *is drawn on a graph, with amplitude plotted against time, straight lines and right angles rather than curves appear.*



squib \'skwib\ *n.*

ASTRONAUTICS. A small device that activates a rocket engine by starting the combustion of primer fuel in the igniter.

One kind of SQUIB *resembles a firecracker set off by an electrical current passing through a heat wire.*

stabilizer \'stā-bə-ˌlī-zər\ *n.*

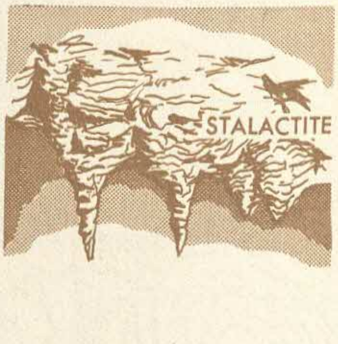
1. AERONAUTICS. A fixed, horizontal tail surface of an aircraft to which elevators are attached; also called horizontal stabilizer.
2. CHEMISTRY. A substance added to a second substance to make the second substance less likely to react; also, a negative catalyst that slows or stops a chemical reaction; *see catalyst.*

A horizontal STABILIZER *may be designed with dihedral to increase an aircraft's longitudinal stability.*

stable equilibrium \'stā-bəl ˌē-kwə-'lib-rē-əm\

PHYSICS. The position to which an object tends to return after being tipped or otherwise moved slightly.

A brick lying on its broadest side is in a state of STABLE EQUILIBRIUM.



stable substance \'stā-bəl 'səb-stən(t)s\

CHEMISTRY. A material or compound that resists decomposition or separation into simpler materials and does not react readily with other substances.

Even a STABLE SUBSTANCE *will decompose at extremely high temperatures, such as those in the sun.*

stalactite \stə-'lak-ˌtīt\ *n.*

EARTH SCIENCE. An icicle-shaped deposit of stone that hangs from the roof of a cave or cavern. Stalactites are formed when

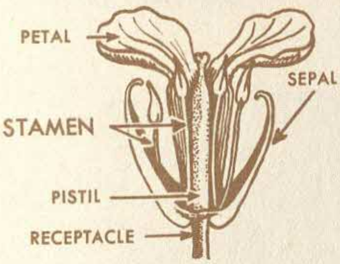
mineral-bearing water seeps through a cave roof and evaporates, leaving the minerals to accumulate in a dripstone deposit; see *dripstone.*

Calcium carbonate from dissolved limestone is the most common mineral that will form a STALACTITE.

stalagmite \stə-'lag-ˌmīt\ *n.*

EARTH SCIENCE. Posts or columns of stone that are deposited on the floor of a cave or cavern. Stalagmites are formed when mineral-bearing water drips onto the floor and evaporates, leaving the minerals to accumulate in a dripstone deposit; see *dripstone.*

A STALAGMITE *and a stalactite sometimes meet to form a column of dripstone from floor to ceiling.*



stamen \'stā-mən\ *n.*

BOTANY. A flower organ consisting of a stalk, or filament, and an anther in which pollen develops.

Flower types that have one or more pistils but no STAMEN *are known as imperfect pistillate flowers.*

staminate flower \'stā-mə-nət 'flau̇(-ə)r\

BOTANY. An imperfect flower bearing stamens but no pistils.

Any monoecious plant has at least one STAMINATE FLOWER *and one distillate flower.*

STAMINATE FLOWER

(WILLOW)

standard deviation \'stan-dərd ˌdē-vē-'ā-shən\

MATHEMATICS. The square root of the average of the squares of numbers that represent differences by which some observations vary from the mean of a series of observations.

The results of a series of experiments may be reduced to statistics and expressed in terms of the mean and STANDARD DEVIATION.

standard solution \'stan-dərd sə-'lü-shən\

CHEMISTRY. A solution of known concentration, usually a water solution for which the concentration of dissolved substance is expressed in terms of molarity or normality. A standard solution is used in chemical analysis as a basis for comparing other solutions.

Ten milliliters of a 1-normal STANDARD SOLUTION *of acid will neutralize exactly 10 milliliters of a 1-normal standard base, or 5 milliliters of a 2-normal base solution.*

standard temperature and pressure \'stan-dərd 'tem-pər-chu̇(ə)r and 'presh-ər\

CHEMISTRY and PHYSICS. A temperature of 0°C. and a pressure of 760 mm. of mercury; abbreviated STP.

The weight of oxygen varies with temperature and pressure, so scientists refer to its weight at STANDARD TEMPERATURE AND PRESSURE, *which is 1.429 g per liter.*

standard time \'stan-dərd 'tīm\

EARTH SCIENCE. A worldwide system of time, based on 24 longitudinal time belts, each 15 degrees wide. The time changes by one hour from one zone to the next, and the time-belt boundaries are varied to meet local needs.

In the United States, Pacific STANDARD TIME *is three hours later than Eastern standard time.*

STANDARD TIME

star \'stär\ *n.*

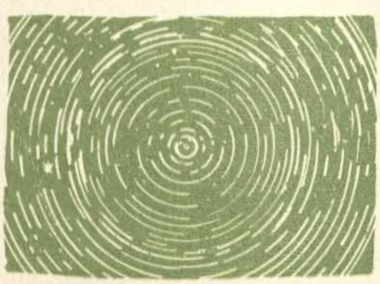
ASTRONOMY. A usually self-luminous sphere of gas in which nuclear reactions produce electromagnetic radiations; see *binary stars, dwarf star, giant star, variable star* and *spectral classes of stars.*

With the exception of the sun, the nearest bright STAR *is Alpha Centauri, about 4⅓ light-years away.*

starch \'stärch\ *n.*

BIOLOGY and CHEMISTRY. Any of a group of polysaccharide carbohydrate compounds present mainly in seeds and tubers but also present as small granules in other plant parts. Starch has molecular weights of approximately 32,000 $(C_6H_{10}O_5)_n$. It is insoluble in water but forms a colloidal solution in hot water. It can be converted to glucose by the chemical addition of water (hydrolysis); see *cellulose.*

Any enzyme (salivary amylaze) in saliva promotes the breakdown of STARCH *to sugar and may be demonstrated by placing some dry cornstarch on the tongue, moistening it with saliva and noting the sweetish taste that occurs after several minutes.*



STAR TRAIL

Time exposure showing star trails around the north star

star stone \'stär 'stōn\

EARTH SCIENCE. A mineral, such as star sapphire, which, because of its internal hexagonal crystal structure, reflects light in a six-rayed, starlike pattern, a phenomenon known as asterism.

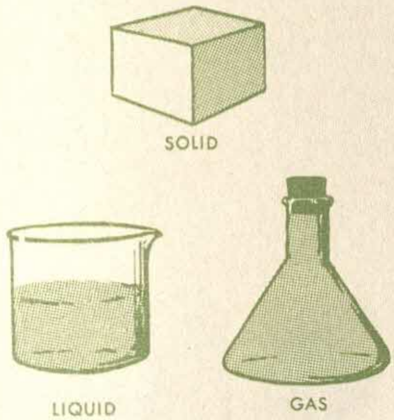
The star ruby is a valuable kind of STAR STONE.

star trail \'stär 'trāl\

ASTRONOMY. A continuous streak or line on a photograph of the night sky. It is produced by the light from a star during a time

exposure by a stationary camera or by a camera moving in a direction different from a star's apparent motion.

If a long-time exposure is made with a camera pointed at Polaris, the North Star, the light from each of the surrounding stars will produce a curved STAR TRAIL.



STATES OF MATTER

states of matter \'stāts əv 'mat-ər\

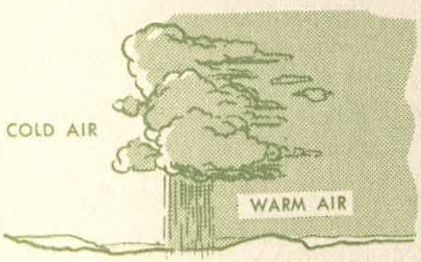
PHYSICS. The physical forms (solid, liquid and gas) in which matter may exist. These states are determined by the separation and rate of motion of the molecules making up the substance. With temperature increase, both molecular separation and rate of motion increase, and, with sufficient temperature, the substance changes from solid to liquid to gas.

Of the three STATES OF MATTER, *only solids have a definite shape.*

static \'stat-ik\ *n.*

PHYSICS. The irregular disturbances and interference received by a radio receiver, usually caused by such electrical phenomena as lightning.

FM radio reception is more free of STATIC *than is AM radio reception.*



STATIONARY FRONT

static electricity \'stat-ik i-ˌlek-'tris-ət-ē\

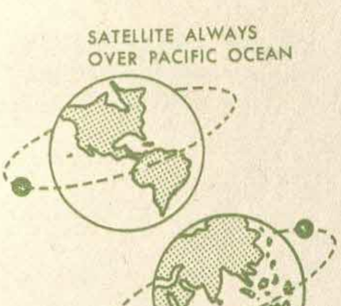
PHYSICS. An electric charge on a nonconducting material, such as rubber or glass; generally, any electric charge that is not in motion, as distinguished from current electricity. Static electricity is frequently produced when one substance is rubbed against another; see *electron.*

STATIC ELECTRICITY *usually disappears (is conducted away) from a charged object faster on a humid day than on a dry day.*

stationary front \'stā-shə-ˌner-ē 'frənt\

EARTH SCIENCE. The boundary, or line, between two air masses that remain in one place for a period of time.

A STATIONARY FRONT *may cause rainfall in one area for several days.*



STATIONARY ORBIT

stationary orbit \'stā-shə-ˌner-ē 'ȯr-bət\

ASTRONAUTICS. The path of a satellite around a celestial body in a plane perpendicular to the axis of rotation of the celestial body, in the same direction as the rotation of the celestial body and at such a distance from the celestial body that the satellite makes one revolution while the celestial body makes one rotation.

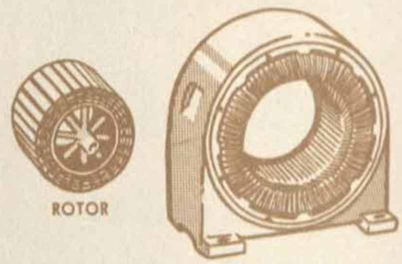
An artificial satellite in STATIONARY ORBIT *constantly maintains*

the same position relative to an observer on earth and, therefore, has tremendous potential in space communication.

stator \'stāt-ər\ *n.*

ENGINEERING and PHYSICS. The stationary coil or magnet of an electrical generator. It may be a coil of wire in which electric current is generated (induced), or the coil producing the magnetic field that induces current; see *rotor.*

If the STATOR *of an electrical generator carries induced current, it is called an armature.*



STATOR

steady state universe \'sted-ē 'stāt 'yü-nə-ˌvərs\

ASTRONOMY and PHYSICS. A theory stating that the universe is in a condition of dynamic equilibrium in which processes that change matter to energy in stars are balanced by processes that change energy into matter in other regions of space. It may be thought of as a matter-energy cycle in which (*a*) thinly-spread particles accumulate in a region of space, and (*b*) the particles are compressed together by mutual gravitational attraction to form stars in which (*c*) nuclear processes are started, transforming matter into energy that is radiated into space and (*d*) that eventually changes into particles, thus completing the cycle.

Acceptance of the theory of a STEADY STATE UNIVERSE *leads to the prediction that galaxies should be more or less uniformly distributed throughout the universe.*

stearic acid \stē-'ar-ik 'as-əd\

CHEMISTRY. A white powder, insoluble in water and occurring in many natural fats as an ester (glyceryl tristearate). It is an organic acid; see *organic acid* and *ester.*

STEARIC ACID *will react with a solution of lye (sodium hydroxide) to form a soap.*

STEARIC ACID — $C_{17}H_{35}COOH$

steel \'stēl\ *n.*

CHEMISTRY. Nearly-pure iron in which a small amount (0.3 percent to 1.7 percent) of carbon is dissolved. It may contain other elements, such as sulfur and phosphorus, that are impurities or nickel and chromium that give the mixture desirable properties; see *alloy.*

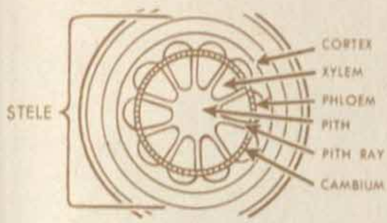
STEEL *that contains 18 percent chromium and 8 percent nickel is classed as stainless steel because it resists corrosion.*

stele \'stē(ə)l\ *n.*

BOTANY. The central, vascular region of a stem or root. A stele

usually contains tissues of xylem, phloem, cambium, rays and pith and is surrounded by cortex.

The vascular bundles that make up the STELE *in the stem of a dicot may be separated by pith rays that appear as spokes in a microscopic cross section.*



STELE

stellar evolution \'stel-ər ˌev-ə-'lü-shən\

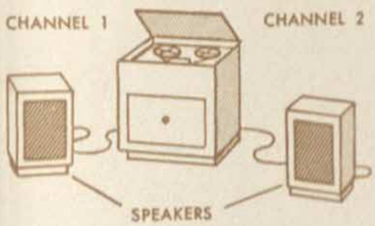
ASTRONOMY. The development or series of changes that stars theoretically undergo over a period of billions of years; see *main-sequence stars, spectral classes of stars* and *spectrum-luminosity diagram.*

According to our present theories of STELLAR EVOLUTION, *the sun, in 5,000 million years, will become much brighter and expand so much that it will engulf several of its nearer planets.*

stereoisomers \ˌster-ē-ō-'ī-sə-mərz\ *n.*

CHEMISTRY. Two molecules that have the same number and kinds of atoms and the same kinds of bonding between atoms but that are arranged differently in space, one being the mirror image of the other; also, a special kind of isomer that is frequently the result of four different groups of atoms being attached to a central carbon atom.

In 1848, Louis Pasteur was able to separate two STEREOISOMERS *from crystals of tartaric acid by using a magnifying glass to detect slight differences in their crystal structure.*



STEREOPHONIC

stereophonic \ˌster-ē-ə-'fän-ik\ *adj.*

ENGINEERING. Referring to a method of reproducing sound, in which two or more microphones are used to record the sound on separate tracks. Each track is then reproduced through a separate loudspeaker, giving an effect of depth.

If one ear is covered, the STEREOPHONIC *effect is decreased.*

stereoscopic \ˌster-ē-ə-'skäp-ik\ *adj.*

PHYSICS. Referring to the kind of vision in which objects, images or scenes are perceived as having depth. Stereoscopic vision is characteristic of all animals that have two eyes pointed forward; see *binocular vision.*

The ability to judge distance depends largely on STEREOSCOPIC *vision.*

STEREOSCOPIC

sterile \'ster-əl\ *adj.*

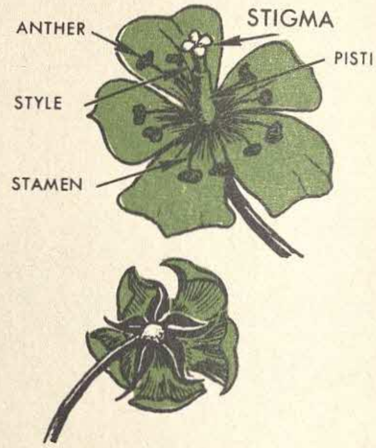
BIOLOGY and MEDICINE. Referring to any object made free of microorganisms by physical (heat) or chemical (antiseptic) means; also, referring to any plant or animal unable to reproduce.

Surgical instruments are usually made STERILE *by very hot steam under pressure.*

sterol \\'sti(ə)r-ˌȯl\\ *n.*

CHEMISTRY. Any one of a group of complex alcohols. A sterol has a high molecular weight and is usually a crystalline solid that is not soluble in water.

The most common STEROL *is cholesterol, found in the human spinal cord, blood and gallstones.*



stigma \\'stig-mə\\ *n.*

BOTANY. The surface at the tip of the style of a pistil in flowers. It is usually covered by sticky fluid or hairs and receives pollen from the stamen.

A pollen grain usually germinates best in a flower of the same or related kind of plant because of the composition of the substances that cover the STIGMA.

stimulus \\'stim-yə-ləs\\ n.

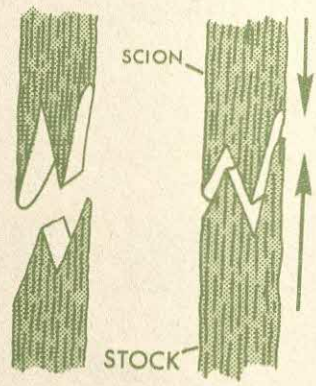
BIOLOGY and PHYSIOLOGY. Any condition or influence, either in the environment or in an organism, that brings about a functional reaction; in animals with nervous systems, a condition that activates a receptor, bringing about a response.

One important characteristic of all living things is their ability to respond to a STIMULUS.

stipule \\'stip-(ˌ)yü(ə)l\\ *n.*

BOTANY. One of a pair of variously-shaped structures that occurs at or near the base of the petiole of a leaf in many kinds of plants.

The leaflike STIPULE *may be present when a leaf is young, but it dies and falls off as the leaf matures.*



stock \\'stäk\\ *n.*

1. BOTANY. The stem or root of a plant in which another stem or bud, the scion, is inserted in grafting. 2. EARTH SCIENCE. A body of igneous rock that has been forced into older rock. A stock usually cuts across sedimentary beds, increases in size downward and is less than 40 square miles in area. It may be a small batholith and is sometimes called a boss.

When the stem of a tomato plant is grafted to the base of a potato plant, the plant sometimes yields potatoes on the STOCK *and tomatoes on the scion.*

stoichiometry \\ˌstȯi-kē-'äm-ə-trē\\ *n.*

CHEMISTRY. The mathematics of chemical reactions. It deals especially with calculations involving the amounts and kinds of

reactants and products. It also deals with the amount of energy absorbed or given off by reactions.

By STOICHIOMETRY, *it is possible to calculate that 18 pounds of water, if decomposed into its elements, would produce about 16 pounds of oxygen and 2 pounds of hydrogen.*

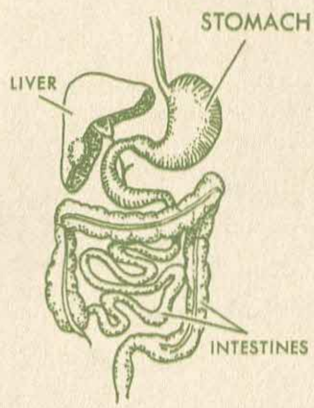
stolon \'stō-lən\ *n.*

BOTANY. A shoot or branch of a stem that extends horizontally at or below ground level and gives rise to upright stems; a runner.

In strawberries, a STOLON *may grow relatively far from the parent plant.*

stomach \'stəm-ək\ *n.*

ANATOMY and ZOOLOGY. The pouchlike digestive organ in man and many other animals.



In man, the inner lining (mucosa) of the STOMACH *contains millions of tubular glands that secrete gastric juice.*

stomates \'stō-ˌmāts\ *n.*

BOTANY. Small openings between pairs of guard cells in the epidermis of a leaf or stem. They regulate the rate of water loss (transpiration) and serve as openings through which the gases involved in photosynthesis and respiration diffuse.

STOMATES *are closed by the shrinking and straightening of the guard cells and are opened by the swelling and arching of these cells.*

storage battery \'stōr-ij 'bat-ə-rē\

CHEMISTRY and PHYSICS. Two or more electrolytic cells connected together to act as a single source of electricity. Such a battery produces direct current electricity from chemical reactions and may also be recharged by direct current electricity.

A common automobile STORAGE BATTERY *contains six 2-volt cells connected in series to produce a total of 12 volts.*

STOMATES

STP

An abbreviation for standard temperature and pressure. See *standard temperature and pressure.*

straight angle \'strāt 'aŋ-gəl\

MATHEMATICS. An angle equal to 180 degrees or π radians.

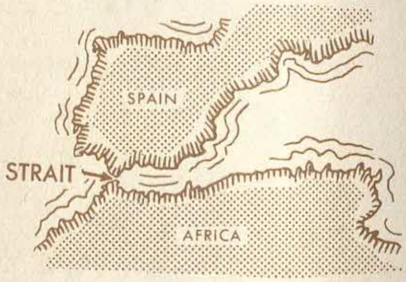
The sides of a STRAIGHT ANGLE *lie on the same straight line.*

straight-chain compounds \'strāt 'chān 'käm-paůndz\
CHEMISTRY. Organic compounds that have molecules containing four or more atoms joined in a continuous line. They are distinguished from branched-chain compounds, in which some of the atoms form side chains, and also from ring compounds, in which four or more atoms are joined in a closed loop.

Gasoline consisting of STRAIGHT-CHAIN COMPOUNDS *has a lower antiknock (octane) rating than gasoline composed of branched-chain compounds.*

strain \'strān\ *n.*
1. PHYSICS. A change of shape produced by such forces (stresses) as pulling (tension), squeezing (compression), bending and twisting. 2. BIOLOGY. A population of plants or animals, within a variety or species, having distinguished physiological or, less commonly, structural characteristics, such as DDT-resistant flies or wilt-resistant tomatoes.

If a given stress is doubled, the amount of STRAIN *will double, providing the elastic limit is not exceeded.*



strait \'strāt\ *n.*
EARTH SCIENCE. A narrow stretch of sea or a waterway connecting two larger bodies of water; see *sound* (3).

A STRAIT *may separate two continents, two islands or an island and a mainland.*

stratified \'strat-ə-ˌfīd\ *adj.*
EARTH SCIENCE. Referring to horizontal layers, beds or strata of rocks, soils or clouds.

Any of the transporting agents, such as wind, water or ice, may produce STRATIFIED *deposits.*

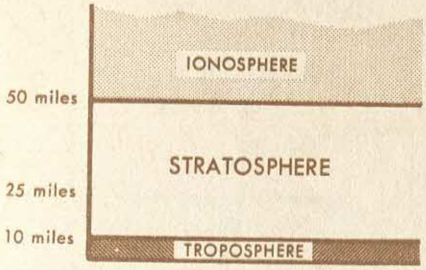
STRATOCUMULUS
(CLOUDS)

stratigraphy \strə-'tig-rə-fē\ *n.*
EARTH SCIENCE. A branch of geology concerned with the formation, arrangement, content and sequence of stratified rock.

The study of STRATIGRAPHY *has revealed much about the history of the earth.*

stratocumulus \ˌstrāt-(ˌ)ō-'kyü-myə-ləs\ *adj.*
EARTH SCIENCE. Referring to a type of low cloud, between the earth's surface and 6,500 feet, composed of globular masses or rolls that often join and form a cloud cover.

STRATOCUMULUS *clouds covering a large part of the sky usually have a wavy appearance.*



stratopause \\'strat-ə-ˌpȯz\\ *n.*

EARTH SCIENCE. A transition layer, or zone, of the earth's atmosphere, extending from the top of the stratosphere to the bottom of the ionosphere.

The STRATOPAUSE *occurs at an altitude of about 50 miles.*

stratosphere \\'strat-ə-ˌsfi(ə)r\\ *n.*

EARTH SCIENCE. A layer of the earth's atmosphere, extending from about 6 to 50 miles above the surface, the heights being less at the poles and greater at the equator.

Clouds rarely form in the STRATOSPHERE, *since there is almost no water vapor or dust present.*

stratum \\'strāt-əm\\ *n.*

EARTH SCIENCE. A bed or layer of sedimentary rock. The rock layers above and below it usually differ in thickness and composition.

A rock STRATUM *may change texture or content if subjected to pressure, water or heat.*

STRATUM

stratus \\'strāt-əs\\ *adj.*

EARTH SCIENCE. Referring to a type of low cloud that occurs in a uniform layer less than 6,500 feet above the earth's surface.

STRATUS *clouds look like fog and give the sky a hazy appearance.*

streak \\'strēk\\ *n.*

EARTH SCIENCE. The color shown when a mineral is powdered or rubbed against unglazed porcelain.

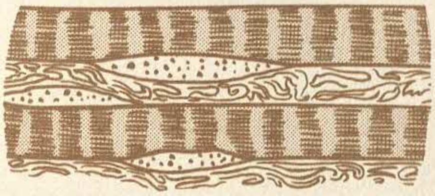
A mineral's STREAK *often differs from its surface color and is a means of identifying some minerals.*

stress \\'stres\\ *n.*

PHYSICS. A force that tends to change the shape of an object. It is frequently expressed as force per unit area, such as pounds per square inch; see *strain* (1).

A piece of paper has low resistance to tearing STRESS *(shear) but has relatively-high resistance to the stress of pulling (tension).*

STRIATED MUSCLES



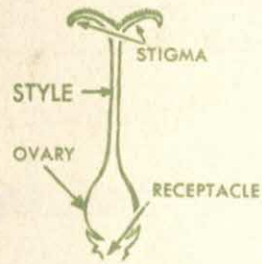
striated muscles \\'strī-ˌāt-əd 'məs-əlz\\

ANATOMY. Muscles characterized by fibers that are divided or marked by microscopic, transverse bands; see *voluntary muscles.*

STRIATED MUSCLES *are usually attached to the bones and move various parts of the body.*

stroboscopic effect \\ˌstrō-bə-ˈskäp-ik i-ˈfekt\\
PHYSICS. An illusion that a moving object is standing still, moving in the opposite direction or moving in jumps instead of moving smoothly. The illusion is caused by viewing the moving object at short intervals rather than continuously or by lighting the moving object with a series of flashes rather than continuously.

The STROBOSCOPIC EFFECT *is often apparent in movies when the spokes of rotating wheels appear to be stationary or moving backward.*

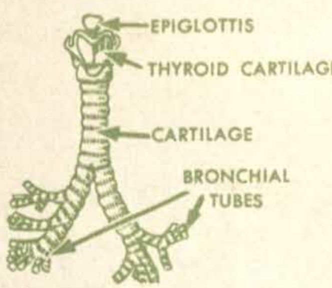


structural formula \\ˈstrək-chə-rəl ˈfȯr-myə-lə\\
CHEMISTRY. A diagram showing the number, kinds and arrangement of atoms in a molecule. It is usually used for covalent compounds in which each covalent bond is represented by a dash (—); see *covalent bond.*

The empirical formula for acetylene is C_2H_2, while its STRUCTURAL FORMULA *is $H—C{\equiv}C—H$.*

style \\ˈstī(ə)l\\ *n.*
BOTANY. The usually-slender part of the pistil of most flowers that extends from the top of the ovary and bears the stigma at its upper end.

Pollen grains germinating on the stigma grow through the middle of the STYLE *to the ovules in the ovary.*



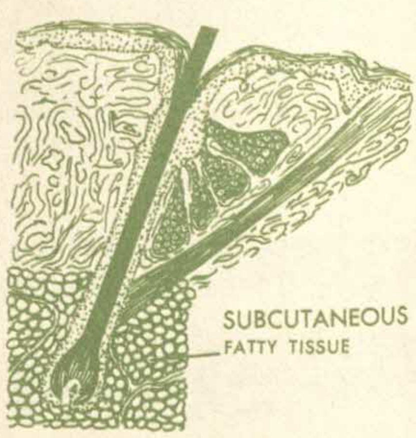
SUBCARTILAGINOUS

subatomic particles \\ˌsəb-ə-ˈtäm-ik ˈpärt-i-kəlz\\
PHYSICS. Particles smaller than atoms, or fragments of atoms that are produced when the nucleus of an atom is disintegrated. More than 30 different types are known, including protons, neutrons, electrons, positrons, antiprotons, mesons and deuterons.

Many SUBATOMIC PARTICLES *have been discovered by observing the disintegration products of radioactive materials with such instruments as Geiger counters and cloud chambers.*

subcartilaginous \\ˌsəb-ˌkärt-ᵊl-ˈaj-ə-nəs\\ *adj.*
ANATOMY and ZOOLOGY. Referring to the area under a cartilage; also, describing a tissue that is partly cartilage.

In man, the larynx, trachea and bronchi contain both cartilaginous and SUBCARTILAGINOUS *tissue.*



subcutaneous \\ˌsəb-kyü-ˈtā-nē-əs\\ *adj.*
ANATOMY and MEDICINE. Referring to the area just beneath the skin.

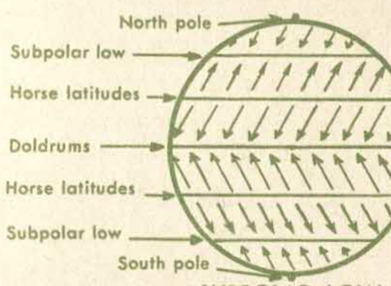
Some hypodermic injections of medicine are injected into SUBCUTANEOUS *tissues.*

subindex \\ˌsəb-ˈin-deks\\ *n.*
MATHEMATICS. A subscript, a subscript of a subscript or a subscript of a superscript.

In the expression 2_{h}, $_h$ is a SUBINDEX.

sublimation \\ˌsəb-lə-ˈmā-shən\\ *n.*
CHEMISTRY and PHYSICS. The change of a solid substance directly into a gas without transition through a noticeable liquid state.

Naphthalene, the main ingredient of mothballs, and dry ice both undergo SUBLIMATION *at room temperature.*



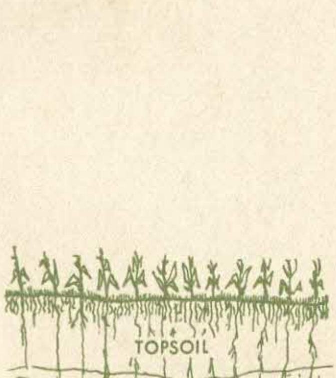
SUBPOLAR LOW

subpolar low \\ˌsəb-ˈpō-lər ˈlō\\
EARTH SCIENCE. Either of the more or less permanent low pressure belts at about 60 degrees north and south latitudes.

The SUBPOLAR LOW *moves toward the equator in the winter season and toward the pole in the summer season.*

subscript \\ˈsəb-ˌskript\\ *n.*
MATHEMATICS. A number or letter written to the right of, and slightly below, a mathematical symbol to distinguish it from other symbols of the same kind or class.

In the polynominal $a_0x^3 + a_1x^2 + a_2x + a_3$, the SUBSCRIPT $_3$ *is associated with the constant term, while the subscripts $_0$, $_1$ and $_2$ are used to distinguish between the various coefficients of the powers of the variable, x.*



SUBSOIL

subset \\ˈsəb-ˌset\\ *n.*
MATHEMATICS. A set all of whose elements are elements of another set.

Set A is a SUBSET *of set B, provided that every element of A is also an element of B.*

subsoil \\ˈsəb-ˌsȯil\\ *n.*
EARTH SCIENCE. The layer of earthy material between the soil and the bedrock. It contains little or no organic matter.

SUBSOIL *may be exposed by soil erosion.*

subsonic \\ˌsəb-ˈsän-ik\\ *adj.*
AERONAUTICS. Referring to a speed less than the speed of sound.

Propeller-driven aircraft travel at SUBSONIC *speeds, while some jets achieve supersonic speeds.*

ABBREVIATIONS

A	ampere
Å	Ångstrom unit
abs	absolute
a-c	alternating current (as an adjective)
amu	atomic mass unit
atm	atmosphere
at. wt	atomic weight
AU	astronomical unit
avdp	avoirdupois
Bev	one billion electron volts
bhp	brake horsepower
bhp-hr	brake horsepower-hour
bp	boiling point
Btu	British thermal unit
C	temperature Celsius; temperature Centigrade
c	candle
cal	calorie
cfm	cubic feet per minute
cfs	cubic feet per second
cgs	centimeter-gram-second (system)
cl	centiliter
cm	centimeter
cm^2	square centimeter
cm^3	cubic centimeter
coef	coefficient
colog	cologarithm
cos	cosine
cot	cotangent
cp	candlepower
csc	cosecant
cu	cubic
cu ft	cubic foot
db	decibel
d-c	direct current (as an adjective)
doz	dozen
E	electromotive force
e	the base of the system of natural logarithms
ev	electron volt
F	temperature Fahrenheit
fp	freezing point
fpm	feet per minute
fps	feet per second
ft	foot; feet
ft^2	square foot
ft^3	cubic foot
ft-c	footcandle
ft-lb	foot-pound
G	universal gravitational constant
g	gram
gal	gallon
g-cal	gram-calorie
gpm	gallons per minute
gps	gallons per second
hr	hour
$h\nu$	photon energy
hp	horsepower
Hz	hertz (cycles per second)
I	electric current
ID	inside diameter
in.	inch
$in.^2$	square inch
$in.^3$	cubic inch
in.-lb	inch-pound
ips	inches per second
j	joule
K	temperature Kelvin (absolute)
kcal	kilocalorie
kg	kilogram
kg-cal	kilogram-calorie
kg-m	kilogram-meter
kg/m^3	kilograms per cubic meter
kgps	kilograms per second
km	kilometer
kv	kilovolt
kw	kilowatt
kwhr	kilowatt-hour
l	liter; lumen
lat	latitude
lb	pound
lb-ft	pound-foot
lb/ft^2	pounds per square foot
lb/ft^3	pounds per cubic foot
lb-in.	pound-inch
l-hr	lumen-hour
lin ft	linear foot
log	logarithm (common)
$\log_e$	logarithm (natural)
long.	longitude
m	meter; minute (time, in astronomical circles)
m^2	square meter
m^3	cubic meter
ma	milliampere
Mev	one million electron volts
mg	milligram
mh	millihenry
mi	mile
mi^2	square mile
min	minute
m-kg	meter-kilogram
ml	milliliter
mm	millimeter
mm^2	square millimeter
mm^3	cubic millimeter
$m\mu$	millimicron
mph	miles per hour
mphps	miles per hour per second
mv	millivolt
N	Avogadro's constant
$n!$	factorial n
OD	outside diameter
oz	ounce
pH	rating on acid-alkaline scale
ppm	parts per million
psi	pounds per square inch
psia	pounds per square inch absolute
R	temperature Reaumur; resistance
RA	right ascension
rpm	revolutions per minute
rps	revolutions per second
sec	secant; second
sin	sine
sp gr	specific gravity
sq	square
tan	tangent
V	volt
VA	volt-ampere
W	watt; work
yd	yard
yd^2	square yard
yd^3	cubic yard

SCIENTIFIC SYMBOLS AND ABBREVIATIONS

α	alpha particle
β; β^-	beta particle
β^+	positron
γ	gamma radiation
Δ	a small change; heat
λ	wavelength; radioactive-decay constant
ma	milliampere
μc	microcurie
μf	microfarad
μin.	microinch
μm	micron
$\mu\mu$	micromicron
$\mu\mu$f	micromicrofarad
ν	frequency; neutrino
π	3.14159; osmotic pressure
Σ	the sum of
σ	nuclear cross section (barns); area
Ω	electrical resistance (ohms)
ω	angular speed; angular velocity
$'$	minute (angular measure)
$''$	second (angular measure)
♂	male
♀	female
$>$	is greater than
$<$	is less than
$\propto$	is proportional to
∞	infinity
$\sqrt{\ }$	square root of
$^\circ$	degrees; temperature; angle measurement (example, 30°)
[ ]	molar concentration
$+$	positive electric charge; mixed with; plus
$-$	negative electric charge; single covalent bond; minus
$=$	equals; double covalent bond; produces
$\neq$	does not equal
$\equiv$	triple covalent bond
$\rightarrow$	produces; forms; chemical reaction
$\rightleftarrows$	reversible chemical reaction
$\uparrow$	gas produced by a chemical reaction
$\downarrow$	precipitate produced by a chemical reaction
$^\bullet$	radioactive substance (follows symbol of element; example, $Cl^\bullet$)